ISSN 2686-9373

ВЕСТНИК СОВРЕМЕННЫХ ЦИФРОВЫХ ТЕХНОЛОГИЙ

24. 2025 (СЕНТЯБРЬ)

Издание включено в перечень ВАК (специальности: 2.3.2, 2.3.6, 2.3.8, 5.2.4)

ВЕСТНИК

СОВРЕМЕННЫХ ЦИФРОВЫХ ТЕХНОЛОГИЙ

НАУЧНО-ПРАКТИЧЕСКИЙ ЖУРНАЛ

№24
СЕНТЯБРЬ 2025

www.c3da.org

ISSN 2686-9373

Издатели: *Российский государственный социальный университет*
Ассоциация РКЦФА

Адрес редакции и издателя: 129226, Москва,
ул. Вильгельма Пика, д.4, стр.1
www.c3da.org

Подписано в печать 25.09.2025 г.
Тираж 500 экз.
Подписной индекс в каталоге «Пресса России»: 79111

Свидетельство о регистрации СМИ
ПИ № ФС 77-76187 от 08.07.2019 г.

Журнал включен в перечень рецензируемых научных изданий ВАК, в которых должны быть опубликованы основные результаты диссертаций на соискание ученой степени кандидата наук, на соискание ученой степени доктора наук.
(2.3.2) Вычислительные системы и их элементы
(2.3.6) Методы и системы защиты информации, информационная безопасность
(2.3.8) Информатика и информационные процессы
(5.2.4) Финансы

РЕДАКЦИОННЫЙ СОВЕТ

РЕДАКЦИОННОЕ ПРИМЕЧАНИЕ

Материалы двадцать четвертого выпуска нашего журнала в значительной степени посвящены практической реализации и концептуальным вопросам разработки технологий искусственного интеллекта и обеспечения информационной безопасности и отражают актуальные тренды развития доверенных и надежных систем.

Выпуск открывает концептуально значимая работа главного редактора журнала Андрея Щербакова **«Практическая модель искусственного сознания»**, дополняющая новыми положениями парадигму семантического искусственного интеллекта. В статье рассматривается общий аспект моделирования сознания на основе динамических когнитивных процессов. На базе теоретико-множественного подхода сформулированы общие требования к модели сознания, показана ее оптимальность, включая исходную мультиязыковость и отсутствие галлюцинаций. Предложенная модель решает важную задачу противодействия генерации ложного контента, характерную для существующих языковых моделей. Ученым-практикам и техническим специалистам могут быть полезны фрагменты кода, описывающие отдельные технические аспекты моделирования процессов сознания.

Статья **«Концептуализация разработки человекоориентированного искусственного интеллекта: междисциплинарный подход»** Владимира Свинаренко посвящена формированию комплексного взгляда на человекоориентированный искусственный интеллект, формулированию его основных характеристик и демонстрации преимуществ междисциплинарной парадигмы по сравнению с традиционными, техноцентричными подходами. Предмет исследования определяется автором как система, которая не только выполняет технические задачи, но и соответствует этическим нормам, является понятной, контролируемой и надежной для пользователя. Важным для разработки и реализации технологий ИИ является вывод о том, что формирование парадигмы человекоориентированного ИИ невозможно при использовании потенциала одной дисциплины.

В рамках следующей работы **«Эволюция культуры информационной безопасности: организационно-технические и социокультурные факторы»** Павла Былевского исследовано развитие культуры информационной безопасности под влиянием организационно-технических и социально-культурных факторов. Отмечается, что значение социально-культурных факторов достигает максимума в результате цифровой трансформации. Рассматриваются основные вехи развития узкопрофессиональной деятельности по защите корпоративной тайны до общегражданской культуры информационной безопасности повседневного быта, традиционных ценностей и культурной идентичности. Статья будет интересна и полезна широкому кругу читателей.

В статье **«Современные методы менеджмента информационной безопасности, основанные на риск-ориентированном подходе»** Ильи Белошицкого и Марины Толстых приведен обзор современных подходов к управлению киберрисками с учетом реалий в сфере защиты информации, описаны методологические основы, преимущества и ограничения основных подходов. Особое внимание уделяется современным киберугрозам, актуальным для российских организаций, и новым угрозам, связанным с применением технологий искусственного интеллекта. Также авторами представлены практические рекомендации для повышения киберустойчивости организации.

Статья Сергея Мирзояна **«Статистические методы оценки криптостойкости генераторов псевдослучайных чисел»** главным образом посвящена применению теста хи-квадрат для анализа случайности выходных последовательностей и формализации понятия криптостойкости как проверяемой статистической гипотезы. В ней представлены результаты экспериментального анализа различных типов ГПСЧ с использованием предложенного подхода и показана эффективность теста хи-квадрат в выявлении скрытых зависимостей между элементами последовательности, что позволяет использовать его в качестве инструмента оценки криптографической надежности и свойств генераторов случайных чисел.

В работе **«Математическая модель для оценки уровня зрелости системы кибербезопасности организации»** тандема авторов рассматриваются вопросы определения состояния информационной безопасности организации на основе показателей, прямо или опосредованно связанных с деятельностью субъекта. Приводится математическая модель для количественной оценки зрелости информационной безопасности организации. Авторами получена зависимость уровня зрелости от различных факторов, прежде всего ресурсов, внутренних и внешних факторов, уровня подготовки персонала подразделений, ответственных за данное направление.

В статье **«Симметрия в криптографии»** Игоря Кириллова обсуждается общее понятие и варианты определений симметрии в конкретных областях науки. Общее определение трактует симметрию прежде всего как философскую категорию и практически не отражает сути явления с позиции математики. В связи с этим автор рассматривает некоторые принципиальные вопросы определения понятия симметрии в криптографии и формулирует неточности его использования в отдельных криптографических публикациях.

В новом произведении **«Дневник доцента Ковалёва»** наш постоянный автор – писатель Егор Федоров – расскажет читательской аудитории журнала поучительную историю от лица ученого-практика в области нейробиологии, исследующего влияние чипирования на поведенческие паттерны подопытного. В качестве основной идеи явственно проступает угроза утраты смысла всякой изначально осознанной деятельности, а затем – и смысла жизни, при замещении человеческой способности к проявлению эмоций только логическими познавательными функциями.

СОДЕРЖАНИЕ

УДК: 004.8

Практическая модель искусственного сознания

A.Yu. Shcherbakov

Practical Model of Artificial Consciousness

Abstract. *The article considers the general aspect of consciousness modeling based on dynamic cognitive processes. According to the set-theoretical approach, general requirements for the consciousness model are formulated, its optimality is shown, including the initial multilingualism. Code fragments describing individual technical aspects of consciousness process modeling are provided. The proposed model solves the important problem of false content generation, which is typical for existing language models.*

Keywords: *artificial consciousness, artificial intelligence, cognitions, cognitive system, semantic artificial intelligence.*

А.Ю. Щербаков
Доктор технических наук, профессор, заведующий кафедрой когнитивно-аналитических и нейро-прикладных технологий Российского государственного социального университета, ведущий научный сотрудник Государственного университета управления.
E-mail: x509@ras.ru

Аннотация. *В статье рассматривается общий аспект моделирования сознания на основе динамических когнитивных процессов. В соответствии с теоретико-множественным подходом сформулированы общие требования к модели сознания, показана ее оптимальность, включая исходную мультиязыковость. Приводятся фрагменты кода, описывающие отдельные технические аспекты моделирования процессов сознания. Предложенная модель решает важную задачу противодействия генерации ложного контента, характерной для существующих языковых моделей.*

Ключевые слова: *искусственное сознание, искусственный интеллект, когниции, когнитивная система, семантический искусственный интеллект.*

ТЕОРЕТИКО-МНОЖЕСТВЕННАЯ МОДЕЛЬ ИСКУССТВЕННОГО СОЗНАНИЯ

Большие языковые модели в настоящее время вполне успешно имитируют «творческие» процессы, такие как формирование различного рода целевых текстов (бизнес-планов, презентаций, рефератов) и процессы разработки программного кода, по сути заменяя индивидуального творца-человека на некоторый «обобщенный средний опыт», полученный в результате закрытого обучения нейросети некоторым неизвестным корпусом текстов [1].

При этом вполне очевидно, что для перспективного развития существующие языковые модели непригодны, в первую очередь из-за их непрозрачности и весьма вероятных сценариев «галлюцинирования», которые отмечаются во многих источниках.

В работе [2] была предложена конструктивная модель для прототипов семантического искусственного интеллекта (ИИ), реализующих концепт семантического мышления, которая опирается на динамическое взаимодействие трех множеств: T, R и Z (рис. 1).

При этом образуется совокупность следующих объектов, которые представлены в виде множеств:

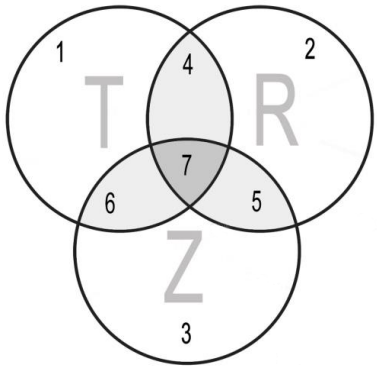


Рис. 1. Модель множеств для семантического искусственного интеллекта [2]

1 – объекты, входящие только в T,
2 – объекты, входящие только в R,
3 – объекты, входящие только в Z,
4 – пересечение множеств T и R,
5 – пересечение множеств R и Z,
6 – пересечение множеств T и Z,
7 – пересечение множеств T, R и Z.

Эта модель позволяет внести в процесс функционирования ИИ как зависимость от дискретного времени, так и архитектурные особенности, связанные с объектами рассмотрения (окружающий мир,

сознание и подсознание), либо категории самого субъективного времени (прошлое, настоящее и будущее).

Пусть в текущий момент дискретного времени T – это информация от окружающего мира, выраженная в виде понятий (когниций), R – сознание мыслящей (когнитивной) системы (КоС) или субъекта коммуникаций в виде некоторого фиксированного на данный момент дискретного времени набора осмысленных (имеющихся в распоряжении КоС) понятий (слов, объектов), Z – подсознание (понимаемое как область интуитивных восприятий), также выраженное в понятиях (словах, выражениях).

В этой модели заложена как внешняя (с окружающим миром, как источником информации, и с внешним субъектом), так и внутренняя коммуникативность.

Тогда возможны следующие интерпретации множеств:

1 – информация окружающего мира, не воспринятая когнитивной системой (текущее непознаваемое или границы опыта),

2 – набор понятий сознания, не участвующих в процессах восприятия (пассивные знания и навыки),

3 – подсознательная область, пассивная на данный момент,

4 – область соприкосновения окружающего мира и сознания (например, вся совокупность образов, поданная сознанию или воспринятая сознанием от органов чувств),

5 – область влияния бессознательного в сознании,

6 – область интерпретации подсознанием окружающего мира,

7 – точка текущего восприятия (мысль, точка сборки или фокус сознания) КоС.

Следует учитывать и то, что такая модель сознания не противоречит популярному на сегодняшний день мнению о том, что мысли (множество 7) зарождаются вне КоС (в первую очередь, человека).

Здесь необходимо отметить некоторую «тавтологичность»: мы называем «сознанием» и всю модель, и его часть (R), ответственную за принятие «логичных» решений. Имеются в виду такие решения, которые одинаково интерпретируются всеми остальными носителями сознания, находящимися вне модели, то есть коммуницирующими с выделенным для рассмотрения отдельным сознанием через внешний мир (T).

Таким образом, с одной стороны индивидуальность сознания может быть помещена в область подсознательного, либо изменение элементов множества R может также содержать индивидуальные функции или операции.

Расширим рассмотрение тем, что допустим, что какое-то из множеств может быть пустым (условно обозначим ситуацию как «0») и непустым (обозначим как «1»). Тогда мы получим 8 квазистатических состояний сознания, которые можно описать с точки зрения человеческих аналогий в таблице 1.

Таблица 1

Состояния сознания

№	T	R	Z	Описание
1	0	0	0	Условное «небытие»
2	0	0	1	Активно только подсознание (условный «сон разума»)
3	0	1	0	Осознанное состояние без информации от внешнего мира (мышление с «закрытыми глазами»)
4	0	1	1	Сон (активны сознание и подсознание, сигналов от внешнего мира нет)
5	1	0	0	Объективный внешний мир при отсутствии наблюдателя
6	1	0	1	Подсознательное восприятие внешнего мира (сознание выключено)
7	1	1	0	Рациональное восприятие ОМ
8	1	1	1	Обычное (повседневное) восприятие

МОДЕЛЬ СЕМИ МНОЖЕСТВ И КОГНИЦИИ

Попытаемся гармонизировать и пояснить сформулированную теоретико-множественную модель с учетом уже устоявшихся терминов.

В современной теории и практике когниции – это процессы познания, включающие в себя восприятие, мышление, память, внимание, речь и другие умственные операции, необходимые для обработки информации и формирования знаний о мире. Иначе говоря, когниции – это всё, что относится к человеческому мышлению и уму.

В первую очередь когниции можно рассматривать как познавательные процессы, включающие в себя восприятие, внимание, память, мышление, речь, решение проблем и другие процессы, которые позволяют носителю сознания обрабатывать информацию и понимать окружающий мир.

Исходя из этого модель когнитивной системы должна реализовывать потоки информации между описанными множествами, такие как поток между множеством T и R, описывающий процессы познания. Например, можно моделировать процесс чтения, когда потоком является последовательность слов. Напомним также, что поток информации рассматривается от одного объекта к другому при помощи активной сущности – субъекта.

Другой стороной когниций являются знания, убеждения, установки и представления о мире, которые формируются на основе опыта и информации, полученной из окружающего мира, сознания и подсознания (категорический нравственный императив).

Соответственно, модель КоС должна наполнять множества R и Z понятиями, коррелирующими с индивидуальностью носителя сознания. Поскольку мы рассматриваем субъектов как операторы множеств T, R и Z, нам важно отделить субъектов, выполняющих «операции» сознания, от его носителя – системы, в которой сознание в целом функционально реализовано.

Когниции играют важную роль и в формировании поведения, так как мысли (множество 7 в модели) и убеждения влияют на ассоциации, решения и действия.

Таким образом, субъекты, изменяющие множества R и Z должны иметь возможность изменений функционирования, например, за счет активизации новых субъектов, оперирующих указанными множествами.

Итак, мы зафиксировали, что когнитивная деятельность связана с процессом ***познания***, который мы понимаем как деятельность субъекта (в частности, человека), направленную на получение нового знания об окружающем мире. Эта деятельность включает в себя отражение объективной реальности (T) в сознании и использование различных методов и средств (внутренних функций модели) для достижения этой цели.

Процесс познания предполагает взаимодействие между субъектом, который познает, и объектом, который познается, с применением различных средств (органов чувств, мышления, языка, инструментов и методов), и может включать в себя различные этапы, такие как восприятие, формирование понятий, выдвижение гипотез и их проверка.

Познание может быть чувственным (основанным на ощущениях) или рациональным (основанным на мышлении), может принимать различные формы, включая обыденное, научное познание, искусство и религию (традиции). При этом основной целью познания является получение нового знания, которое является отражением объективной реальности и может быть использовано для понимания мира и принятия решений. «От живого созерцания к абстрактному мышлению и от него к практике – таков диалектический путь познания истины…», как писал В.И.Ленин[1].

Очевидно, что в целом познание – это сложный и многогранный процесс, который является неотъемлемой частью жизни человека и общества. Поэтому существует ряд специальных наук и научных дисциплин, исследующих этот предмет: когнитивная психология, научная методология, история науки, науковедение, социология знания и др. Однако большинство этих наук изучает познание, рассматривая только его отдельные аспекты. В целом познание остаётся особым предметом изучения философии.

Отдельный тип целостного познания мира представляет собой философское познание, особенностью которого является стремление выйти за пределы фрагментарной действительности и найти фундаментальные принципы и основы бытия, определить место человека в нём. Философское познание основано на определённых мировоззренческих предпосылках. В процессе философского познания субъект стремится не только понять бытие и место человека в нём, но и показать, какими они должны быть (аксиология), то есть стремится создать идеал, содержание которого будет обусловлено избранными философом мировоззренческими постулатами.

[1] Ленин В.И. Философские тетради. – Полн. собр. соч., т. 29, с. 152-153.

Научное познание, в отличие от других многообразных форм познания — процесс получения объективного знания, отражающего закономерности действительности (реальности) – множество T, объективно существующее вне субъекта. Научное познание имеет троякую задачу: описание, объяснение, предсказание процессов и явлений наблюдаемой действительности.

Относительно семантической модели можно утверждать, что формы познания эквивалентны, при этом моделирование процессов познания зависит от практических целей и задач, стоящих перед учеными и разработчиками.

Таким образом, можем видеть, что дальнейшая разработка модели искусственного сознания актуальна с позиции реконструирования информационных процессов различных видов рационального познания, прежде всего – научного.

Исходя из рассмотренных позиций можно предложить дополнения модели сознания [2].

НЕОБХОДИМЫЕ ДОПОЛНЕНИЯ К МОДЕЛИ

К модели сознания (модели когнитивной системы, КоС) целесообразно сделать следующие важные дополнения:

1. Множества T, R и Z представляют собой набор понятий, описывающих общие для всех носителей сознания одного вида категории и явления, т.е. они являются словами естественного языка.

2. Модель КоС должна наполнять множества R и Z понятиями, коррелирующими с индивидуальностью носителя сознания (поскольку мы рассматриваем субъектов как операторов множеств T, R и Z, нам важно отделить субъектов, выполняющих «операции» сознания, от его носителя – системы в которой сознание в целом функционально реализовано), для моделирования мы можем задавать множества и их мощность случайным образом.

3. С точки зрения моделирования и других аспектов технических и программных реализаций целесообразно вместо слов оперировать токенами – цифровыми образами слов.

4. Модель не должна зависеть от языка и должна быть изначально мультиязыковой, что принципиально не только отличает, но и превосходит существующие БЯМ. Эта позиция принципиально связана с универсальным механизмом токенизации (см.п.3). кроме того, этот принцип даёт возможность существенного снижения вычислительной трудоемкости реализации модели искусственного сознания.

5. В функциях преобразования множеств целесообразно добавить «коэффициент внимания» – с какой вероятностью понятия (слова или их токены) переходят (возможно, с трансформацией) из одного множества в другое? Если для перехода из множества T в множество R эта вероятность действительно является «вниманием», то для обратного перехода – внесения в окружающий мир когниций в виде текстов необходимо оперировать коэффициентом «творческой активности» носителя сознания.

КОМПАКТНАЯ МОДЕЛЬ СОЗНАНИЯ

Без ограничения общности рассмотрим два множества – окружающий мир (T) и сознание (R).

С точки зрения программиста, предлагается ввести следующие параметры модели искусственного сознания:

Фрагмент кода 1.

```
// Параметры модели ИСо
// Максимальный размер входного потока
#define MAX_IN 100
// Текущий размер входного потока
   int cur_in;
// Сам входной поток
   long info_in[MAX_IN];
// Файл входного потока
   unsigned char in_f[64];
// Длина файла входного потока
   long in_len;
// Максимальный размер сознания в понятиях (словах)
#define MAX_SO 20000
// Текущий размер сознания
   int cur_so;
// Сам массив сознания
   long info_so[MAX_SO];
// Вероятность изменений сознания за счет входного потока
   int v_dso;
// коэффициент внимания – вероятность "захвата" слова сознанием
#define ATT 40
// минимальная длина слова
#define MINLEN 3
```

Поясним введенные параметры с точки зрения описанных выше множеств: параметр cur_so описывает мощность множества R, а массив long info_so[MAX_SO] – само текущее содержание сознания в токенах. Токен в описываемой модели – число, длиной 4 байта (long), полученное в результате детерминированной процедуры.

Особенностями множества T является то, что оно представляет собой максимально возможный набор понятий. Для его создания используется некоторый «большой» текст (в демонстрационном примере – сборник ранних рассказов А.П.Чехова [4]. Этот текст индексируется модулем m_ind [2, 3] и из него составляется максимально возможный набор (словарь) понятий. Входной поток содержит последовательность токенов и моделирует либо зрительно восприятие, либо процесс чтения. В рассматриваемой модели токены множества T «подаются на вход» последовательно.

Функция токенизации слов выглядит следующим образом:

Фрагмент кода 2.

```
#define MAX_WORD 32
int xb(char *wd, unsigned char *x)
{
  int i,j,len,part,ost;
  unsigned char wd1[32];
  len=strlen(wd);
  if(len>MAX_WORD) return(-1);
  part=len/32;
  ost =len%32;
  for(i=0;i<32 ;i++) wd1[i]=0;
  for(i=0;i<ost;i++) wd1[i]=wd[i];

  for(i=0;i<part;i++)
    for(j=0;j<32;j++) wd1[j]=wd1[j]^wd[i*32+j];

  for(i=0; i<8; i++) x[i]=0x88^(char)i;
// Восемь итераций алгоритма шифрования ГОСТ28147-89
  imit_fast((unsigned long *)x,(unsigned long *)wd1);
  return(0);
}
```

Входным значением является слово wd, а выходным – токен x.

Фрагмент кода 3 (используемые во фрагменте 2 функции).

```
void imit_fast(unsigned long *s,unsigned long *k)
{
  void  elem_gost(unsigned long *, unsigned long *,unsigned long *);
  unsigned long  cur;

  elem_gost(s,s+1,k  );
  elem_gost(s+1,s,k+1);
  elem_gost(s,s+1,k+2);
  elem_gost(s+1,s,k+3);
  elem_gost(s,s+1,k+4);
  elem_gost(s+1,s,k+5);
  elem_gost(s,s+1,k+6);
  elem_gost(s+1,s,k+7);

  cur=*s;
  *s=*(s+1);
  *(s+1)=cur;
}
void elem_gost_( unsigned long  *aa, unsigned long  *bb, unsigned long  *key)
{
  static unsigned char  pod[1024]=
  {
    0xE4,0xEA,0xE9,0xE2,0xED,0xE8,0xE0,0xEE,0xE6,0xEB,0xE1,0xEC,0xE7,0xEF,0xE5,0xE3,
    0xB4,0xBA,0xB9,0xB2,0xBD,0xB8,0xB0,0xBE,0xB6,0xBB,0xB1,0xBC,0xB7,0xBF,0xB5,0xB3,
    0x44,0x4A,0x49,0x42,0x4D,0x48,0x40,0x4E,0x46,0x4B,0x41,0x4C,0x47,0x4F,0x45,0x43,
    0xC4,0xCA,0xC9,0xC2,0xCD,0xC8,0xC0,0xCE,0xC6,0xCB,0xC1,0xCC,0xC7,0xCF,0xC5,0xC3,
    0x64,0x6A,0x69,0x62,0x6D,0x68,0x60,0x6E,0x66,0x6B,0x61,0x6C,0x67,0x6F,0x65,0x63,
    0xD4,0xDA,0xD9,0xD2,0xDD,0xD8,0xD0,0xDE,0xD6,0xDB,0xD1,0xDC,0xD7,0xDF,0xD5,0xD3,
    0xF4,0xFA,0xF9,0xF2,0xFD,0xF8,0xF0,0xFE,0xF6,0xFB,0xF1,0xFC,0xF7,0xFF,0xF5,0xF3,
    0xA4,0xAA,0xA9,0xA2,0xAD,0xA8,0xA0,0xAE,0xA6,0xAB,0xA1,0xAC,0xA7,0xAF,0xA5,0xA3,
    0x24,0x2A,0x29,0x22,0x2D,0x28,0x20,0x2E,0x26,0x2B,0x21,0x2C,0x27,0x2F,0x25,0x23,
    0x34,0x3A,0x39,0x32,0x3D,0x38,0x30,0x3E,0x36,0x3B,0x31,0x3C,0x37,0x3F,0x35,0x33,
```

```
    0x84,0x8A,0x89,0x82,0x8D,0x88,0x80,0x8E,0x86,0x8B,0x81,0x8C,0x87,0x8F,0x85,0x83,
    0x14,0x1A,0x19,0x12,0x1D,0x18,0x10,0x1E,0x16,0x1B,0x11,0x1C,0x17,0x1F,0x15,0x13,
    0x04,0x0A,0x09,0x02,0x0D,0x08,0x00,0x0E,0x06,0x0B,0x01,0x0C,0x07,0x0F,0x05,0x03,
    0x74,0x7A,0x79,0x72,0x7D,0x78,0x70,0x7E,0x76,0x7B,0x71,0x7C,0x77,0x7F,0x75,0x73,
    0x54,0x5A,0x59,0x52,0x5D,0x58,0x50,0x5E,0x56,0x5B,0x51,0x5C,0x57,0x5F,0x55,0x53,
    0x94,0x9A,0x99,0x92,0x9D,0x98,0x90,0x9E,0x96,0x9B,0x91,0x9C,0x97,0x9F,0x95,0x93,
    0x75,0x78,0x71,0x7D,0x7A,0x73,0x74,0x72,0x7E,0x7F,0x7C,0x77,0x76,0x70,0x79,0x7B,
    0xD5,0xD8,0xD1,0xDD,0xDA,0xD3,0xD4,0xD2,0xDE,0xDF,0xDC,0xD7,0xD6,0xD0,0xD9,0xDB,
    0xA5,0xA8,0xA1,0xAD,0xAA,0xA3,0xA4,0xA2,0xAE,0xAF,0xAC,0xA7,0xA6,0xA0,0xA9,0xAB,
    0x15,0x18,0x11,0x1D,0x1A,0x13,0x14,0x12,0x1E,0x1F,0x1C,0x17,0x16,0x10,0x19,0x1B,
    0x05,0x08,0x01,0x0D,0x0A,0x03,0x04,0x02,0x0E,0x0F,0x0C,0x07,0x06,0x00,0x09,0x0B,
    0x85,0x88,0x81,0x8D,0x8A,0x83,0x84,0x82,0x8E,0x8F,0x8C,0x87,0x86,0x80,0x89,0x8B,
    0x95,0x98,0x91,0x9D,0x9A,0x93,0x94,0x92,0x9E,0x9F,0x9C,0x97,0x96,0x90,0x99,0x9B,
    0xF5,0xF8,0xF1,0xFD,0xFA,0xF3,0xF4,0xF2,0xFE,0xFF,0xFC,0xF7,0xF6,0xF0,0xF9,0xFB,
    0xE5,0xE8,0xE1,0xED,0xEA,0xE3,0xE4,0xE2,0xEE,0xEF,0xEC,0xE7,0xE6,0xE0,0xE9,0xEB,
    0x45,0x48,0x41,0x4D,0x4A,0x43,0x44,0x42,0x4E,0x4F,0x4C,0x47,0x46,0x40,0x49,0x4B,
    0x65,0x68,0x61,0x6D,0x6A,0x63,0x64,0x62,0x6E,0x6F,0x6C,0x67,0x66,0x60,0x69,0x6B,
    0xC5,0xC8,0xC1,0xCD,0xCA,0xC3,0xC4,0xC2,0xCE,0xCF,0xCC,0xC7,0xC6,0xC0,0xC9,0xCB,
    0xB5,0xB8,0xB1,0xBD,0xBA,0xB3,0xB4,0xB2,0xBE,0xBF,0xBC,0xB7,0xB6,0xB0,0xB9,0xBB,
    0x25,0x28,0x21,0x2D,0x2A,0x23,0x24,0x22,0x2E,0x2F,0x2C,0x27,0x26,0x20,0x29,0x2B,
    0x55,0x58,0x51,0x5D,0x5A,0x53,0x54,0x52,0x5E,0x5F,0x5C,0x57,0x56,0x50,0x59,0x5B,
    0x35,0x38,0x31,0x3D,0x3A,0x33,0x34,0x32,0x3E,0x3F,0x3C,0x37,0x36,0x30,0x39,0x3B,
    0x46,0x4C,0x47,0x41,0x45,0x4F,0x4D,0x48,0x44,0x4A,0x49,0x4E,0x40,0x43,0x4B,0x42,
    0xB6,0xBC,0xB7,0xB1,0xB5,0xBF,0xBD,0xB8,0xB4,0xBA,0xB9,0xBE,0xB0,0xB3,0xBB,0xB2,
    0xA6,0xAC,0xA7,0xA1,0xA5,0xAF,0xAD,0xA8,0xA4,0xAA,0xA9,0xAE,0xA0,0xA3,0xAB,0xA2,
    0x06,0x0C,0x07,0x01,0x05,0x0F,0x0D,0x08,0x04,0x0A,0x09,0x0E,0x00,0x03,0x0B,0x02,
    0x76,0x7C,0x77,0x71,0x75,0x7F,0x7D,0x78,0x74,0x7A,0x79,0x7E,0x70,0x73,0x7B,0x72,
    0x26,0x2C,0x27,0x21,0x25,0x2F,0x2D,0x28,0x24,0x2A,0x29,0x2E,0x20,0x23,0x2B,0x22,
    0x16,0x1C,0x17,0x11,0x15,0x1F,0x1D,0x18,0x14,0x1A,0x19,0x1E,0x10,0x13,0x1B,0x12,
    0xD6,0xDC,0xD7,0xD1,0xD5,0xDF,0xDD,0xD8,0xD4,0xDA,0xD9,0xDE,0xD0,0xD3,0xDB,0xD2,
    0x36,0x3C,0x37,0x31,0x35,0x3F,0x3D,0x38,0x34,0x3A,0x39,0x3E,0x30,0x33,0x3B,0x32,
    0x66,0x6C,0x67,0x61,0x65,0x6F,0x6D,0x68,0x64,0x6A,0x69,0x6E,0x60,0x63,0x6B,0x62,
    0x86,0x8C,0x87,0x81,0x85,0x8F,0x8D,0x88,0x84,0x8A,0x89,0x8E,0x80,0x83,0x8B,0x82,
    0x56,0x5C,0x57,0x51,0x55,0x5F,0x5D,0x58,0x54,0x5A,0x59,0x5E,0x50,0x53,0x5B,0x52,
    0x96,0x9C,0x97,0x91,0x95,0x9F,0x9D,0x98,0x94,0x9A,0x99,0x9E,0x90,0x93,0x9B,0x92,
    0xC6,0xCC,0xC7,0xC1,0xC5,0xCF,0xCD,0xC8,0xC4,0xCA,0xC9,0xCE,0xC0,0xC3,0xCB,0xC2,
    0xF6,0xFC,0xF7,0xF1,0xF5,0xFF,0xFD,0xF8,0xF4,0xFA,0xF9,0xFE,0xF0,0xF3,0xFB,0xF2,
    0xE6,0xEC,0xE7,0xE1,0xE5,0xEF,0xED,0xE8,0xE4,0xEA,0xE9,0xEE,0xE0,0xE3,0xEB,0xE2,
    0x1D,0x1B,0x14,0x11,0x13,0x1F,0x15,0x19,0x10,0x1A,0x1E,0x17,0x16,0x18,0x12,0x1C,
    0xFD,0xFB,0xF4,0xF1,0xF3,0xFF,0xF5,0xF9,0xF0,0xFA,0xFE,0xF7,0xF6,0xF8,0xF2,0xFC,
    0xDD,0xDB,0xD4,0xD1,0xD3,0xDF,0xD5,0xD9,0xD0,0xDA,0xDE,0xD7,0xD6,0xD8,0xD2,0xDC,
    0x0D,0x0B,0x04,0x01,0x03,0x0F,0x05,0x09,0x00,0x0A,0x0E,0x07,0x06,0x08,0x02,0x0C,
    0x5D,0x5B,0x54,0x51,0x53,0x5F,0x55,0x59,0x50,0x5A,0x5E,0x57,0x56,0x58,0x52,0x5C,
    0x7D,0x7B,0x74,0x71,0x73,0x7F,0x75,0x79,0x70,0x7A,0x7E,0x77,0x76,0x78,0x72,0x7C,
    0xAD,0xAB,0xA4,0xA1,0xA3,0xAF,0xA5,0xA9,0xA0,0xAA,0xAE,0xA7,0xA6,0xA8,0xA2,0xAC,
    0x4D,0x4B,0x44,0x41,0x43,0x4F,0x45,0x49,0x40,0x4A,0x4E,0x47,0x46,0x48,0x42,0x4C,
    0x9D,0x9B,0x94,0x91,0x93,0x9F,0x95,0x99,0x90,0x9A,0x9E,0x97,0x96,0x98,0x92,0x9C,
    0x2D,0x2B,0x24,0x21,0x23,0x2F,0x25,0x29,0x20,0x2A,0x2E,0x27,0x26,0x28,0x22,0x2C,
    0x3D,0x3B,0x34,0x31,0x33,0x3F,0x35,0x39,0x30,0x3A,0x3E,0x37,0x36,0x38,0x32,0x3C,
    0xED,0xEB,0xE4,0xE1,0xE3,0xEF,0xE5,0xE9,0xE0,0xEA,0xEE,0xE7,0xE6,0xE8,0xE2,0xEC,
    0x6D,0x6B,0x64,0x61,0x63,0x6F,0x65,0x69,0x60,0x6A,0x6E,0x67,0x66,0x68,0x62,0x6C,
    0xBD,0xBB,0xB4,0xB1,0xB3,0xBF,0xB5,0xB9,0xB0,0xBA,0xBE,0xB7,0xB6,0xB8,0xB2,0xBC,
    0x8D,0x8B,0x84,0x81,0x83,0x8F,0x85,0x89,0x80,0x8A,0x8E,0x87,0x86,0x88,0x82,0x8C,
    0xCD,0xCB,0xC4,0xC1,0xC3,0xCF,0xC5,0xC9,0xC0,0xCA,0xCE,0xC7,0xC6,0xC8,0xC2,0xCC
};
```

```
    unsigned char r[4];

    *(unsigned long *)r=*aa+*key;

    * r   =*(pod+    * r   );
    *(r+1)=*(pod+256+*(r+1));
    *(r+2)=*(pod+512+*(r+2));
    *(r+3)=*(pod+768+*(r+3));

    *bb^=(*(unsigned long *)r<<11)|((*(unsigned long *)r>>21)&0x7FF);
}
```

Далее происходит «рождение» искусственного сознания – множество R заполняется случайным образом выбранными словами (в виде токенов).

Фрагмент кода 4.

```
// Заполнение сознания понятиями (словами)
   for(i=0;i<cur_so;i++) info_so[i]=0;
   for(i=0;i<cur_so;i++)
     {
      rr=r_l(c_len1/CSV_PAGE);
      ReadCSVitem(CSV_N,rr,w1);
// Запрет заполнения короткими словами
    if(mstrlen(w1,MAX_WORD)<MINLEN) continue;

      for(j=0;j<32;j++) w2[j]=0;
            for(j=0;j<mstrlen(w1,MAX_WORD);j++)
w2[j]=w1[j];
// Вычисление токена
      xb(w2,xx);

      pp=uint8ToUint32(xx);
      info_so[i]=pp;
     }
```

Ядро модели искусственного сознания

Функционирование модели происходит следующим образом. На условный «вход» модели последовательно подаются токены и каждый токен проверяется на наличие в сознании (пересечение T и R). Если токен обнаруживается в сознании, то происходит «мыслительный акт», производящий поиск ассоциации – нескольких токенов из упорядоченного множества T, и все совпадения выводятся в журнал. Практически эта процедура реализована во фрагменте кода 5. Технически такая процедура исключает “галлюцинирование» искусственного сознания, поскольку все ассоциации реально существуют в множестве T и детерминированно могут быть найдены, при этом порождение несуществующей информации невозможно.

Фрагмент кода 5.

```
// Обработка токенов входного потока
   cc=0;ac=0;
   for(i=0;i<in_len/4;i++)
     {
      pp=ReadSPWitem(in_f,i);
      printf("Step = %03d ",i);
      tg=0;
// поиск токена (слова) из входного потока в сознании
      for(j=0;j<cur_so;j++)
        {
// слово найдено в сознании
        if(info_so[j]==pp)
          {
          printf("[%lx] <%d>",pp,cur_so); cc++; tg=1;
// формируется ассоциация
          for(k=0;k<c_len2/4;k++)
            {
             rr=ReadSPWitem(SPW_N,k);
             if(rr==pp)
               {
                kk=k;
                w0[0]='.'; w0[1]=0;
rr1=ReadSPVitem(SPV_N,kk);
rr2=ReadSPVitem(SPV_N,kk+1);
rr3=ReadSPVitem(SPV_N,kk+2);

                ReadCSVitem(CSV_N,rr1,w1);
                for(m=0;m<32;m++) w2[m]=0;
                      for(m=0;m<mstrlen(w1,MAX_
WORD);m++) w2[m]=w1[m];

                ReadCSVitem(CSV_N,rr2,w1);
                for(m=0;m<32;m++) w3[m]=0;
                      for(m=0;m<mstrlen(w1,MAX_
WORD);m++) w3[m]=w1[m];

                ReadCSVitem(CSV_N,rr3,w1);
                for(m=0;m<32;m++) w4[m]=0;
                      for(m=0;m<mstrlen(w1,MAX_
WORD);m++) w4[m]=w1[m];

               AppLogW(LOGNAME,i,w0,w2,w3,w4);
```

```
                    ac++;
                   }
                 }
               }
             }
           }
    //    если слово не найдено в сознании, то оно с заданной вероятностью может "перетечь" в сознание
          if(tg==0)
           {
            if(r_b(ATT)==1)
             {
    //    защита от перетекания короткого слова
              rr1=ReadSPVitem(in_f,i);
              ReadCSVitem(in_f,rr1,w1);
              if(mstrlen(w1,MAX_WORD)>=MINLEN)
              info_so[cur_so]=pp;cur_so++; cur_so=cur_so%MAX_SO;
                                  printf("+");printf("<%d> ",mstrlen(w1,MAX_WORD));}
             }
           }
          printf("\n");
         }
```

Фрагмент кода 6. Вероятность захвата слова

```
long r_l(long mm)
   {

    unsigned long pp;
    NextRandom16(rnd,rnd_1,rnd);
    pp=uint8ToUint32(rnd);
    pp=pp%mm;
    return(pp);
   }
int r_b(long n)
    {
     long t = r_l(100);
     if(t<=n) return(1);
     else    return(0);
    }
```

Во фрагменте можно заметить важную роль датчика случайных чисел – процедура NextRandom16. Динамическая вероятность захвата слова вычисляется в процедуре случайного «бросания» точки (числа) на отрезок заданной длины (меньший или равный входному параметру n).

ПРАКТИЧЕСКИЙ РЕЗУЛЬТАТ ПОСТРОЕНИЯ АССОЦИАЦИЙ

Как было указано выше, в качестве «мира» модели используется [4]. На вход модели подаются слова из рассказа «Дачница», входящего в тот же сборник.

Фрагмент вывода 1. Название входного потока – d, включает 706 слов

Input stream: d.spw

Input stream = 2824 [706 words]

Происходит инициализация множества R размером 8860 слов

Init SO...

Current volume SO = 8860

В квадратных скобках приводится вычисленное значение токена и совпадение на текущем шаге.

Step = 000 [5ae4c2c5] <8860>
Step = 001
Step = 002
Step = 003 [10b639f1] <8860>
Step = 004 +<7>
Step = 005 +<7>
Step = 006
Step = 007 +<7>
Step = 008
Step = 009
Step = 010 +<7>
Step = 011
Step = 012 +<7>
Step = 013
Step = 014 [c347da7f] <8865>
Step = 015
Step = 016 [9b331831] <8865>

На шаге 0 выделено слово «Дачница» (токен [5ae4c2c5]) и все связанные ассоциации (2).

Step 0: ...> [Дачница] > С > женой
Step 0: ...> [Дачница] > Леля > nn

На шаге 3 – слово «хорошенькая» и 10 ассоциаций в «мире» множества T.

Step 3: ...> [хорошенькая] > двадцатилетняя > блондинка
Step 3: ...> [хорошенькая] > девушка > дочь
Step 3: ...> [хорошенькая] > Только > не
Step 3: ...> [хорошенькая] > думал > я
Step 3: ...> [хорошенькая] > сегодня > сказал
Step 3: ...> [хорошенькая] > Оля > вздохнул
Step 3: ...> [хорошенькая] > Почему > же
Step 3: ...> [хорошенькая] > Я > думаю
Step 3: ...> [хорошенькая] > Саша > сказал
Step 3: ...> [хорошенькая] > женская > головка

На шаге 14 – “перекладину» – две ассоциации.

Step 14: ...> [перекладину] > глядит > вдаль
Step 14: ...> [перекладину] > и > полез...

На шаге 16 – «вдаль»

Step 16: ...> [вдаль] > Всё > далекое
Step 16: ...> [вдаль] > но > в

При этом сознание пополнилось пятью новыми когнициями.

ВЫВОДЫ

Предлагаемый подход в виде циклической обработки поступающих и имеющихся в сознании когниций позволяет создать быстродействующую модель искусственного сознания, допускающего обучение информацией из окружающего мира. При этом множество когниций в сознании растет в процессе обучения.

Механизм ассоциирования позволяет находить совпадения и строить тексты, не содержащие ложной информации. Рассмотренная модель может использоваться самостоятельно, в том числе и для моделирования человеческого сознания, либо дополнять и корректировать действующие языковые модели. Учет множества подсознательного (Z) позволит существенно расширить возможности и полноту модели.

СПИСОК ЛИТЕРАТУРЫ

1. Федоров Е. Языковая модель: диалог или монолог? // Вестник современных цифровых технологий. – 2025. – № 23. – с. 42-66.

2. Щербаков А.Ю. Методологические основы и прототип системы семантического искусственного интеллекта // Научно-техническая информация. Сер. 2. Информационные системы и процессы. – 2022. – № 9. – С. 1-6.

3. Кузьменко В.В., Рязанова А.А., Сантьев А.А., Щербаков А.Ю. Семантические алгоритмы как основа создания доверенных систем искусственного интеллекта // Вестник современных цифровых технологий. – 2022. – № 10. – С. 5-10.

4. Третий том полного собрания сочинений А. П. Чехова – рассказы, юморески 1884-1885. Рассказ «Дачница». URL: https://traumlibrary.ru/book/chekhov-pss30-03/chekhov-pss30-03.html#s002003 (Дата обращения: 15.07.2025)

УДК: 004.8, 001.89

Концептуализация разработки человекоориентированного ИИ: междисциплинарный подход

V.A. Svinarenko

Conceptualization of Human-Centered AI Development: Interdisciplinary Approach

Abstract. *The article is devoted to the formation of a comprehensive vision of human-oriented artificial intelligence (HOAI), defining its main characteristics and demonstrating the advantages of the interdisciplinary paradigm compared to traditional, technocentric approaches. The definition of human-oriented AI as a system that not only performs technical tasks, but also enhances human capabilities, complies with ethical standards, is understandable, controllable and reliable for the user is proposed. The fundamental differences between the interdisciplinary approach and the traditional one are demonstrated, emphasizing the need to move from creating biased, unsafe and unacceptable systems to AI that will serve the benefit of humans and correspond to their values. It is concluded that the formation of the HOAI paradigm is impossible within the framework of a single discipline. The essence of the interdisciplinary approach is defined through the description of the contribution of various disciplines and a model for its integration into the life cycle of AI development is proposed.*

Keywords: *human-oriented artificial intelligence, responsible AI, AI ethics, explainable AI, reliable AI, technocentrism.*

В.А. Свинаренко

Аспирант кафедры управления и информатики в технических системах, Московский государственный технологический университет «СТАНКИН».

Email: svi-svi@mail.ru

Аннотация. *Статья посвящена формированию комплексного видения человекоориентированного искусственного интеллекта (ЧОИИ), формулированию его основных характеристик и демонстрации преимуществ междисциплинарной парадигмы по сравнению с традиционными, техноцентричными подходами. Предложено определение человекоориентированного ИИ как системы, которая не только выполняет технические задачи, но и усиливает человеческие возможности, соответствует этическим нормам, является понятной, контролируемой и надежной для пользователя. Продемонстрированы принципиальные отличия междисциплинарного подхода от традиционного, подчеркивающие необходимость перехода от создания предвзятых, небезопасных и неприемлемых для общества систем к ИИ, который будет служить на благо человека и соответствовать его ценностям. Сделан вывод о том, что формирование парадигмы ЧОИИ невозможно в рамках одной дисциплины. Определена суть междисциплинарного подхода через описание вклада различных дисциплин и предложена модель его интеграции в жизненный цикл разработки ЧОИИ.*

Ключевые слова: *человекоориентированный искусственный интеллект, ответственный ИИ, этика ИИ, объяснимый ИИ, надежный ИИ, техноцентризм.*

ВВЕДЕНИЕ

Современный этап развития информационных технологий неразрывно связан с бурным прогрессом в области искусственного интеллекта (ИИ). Технологии машинного обучения, нейронных сетей и обработки естественного языка проникают во все сферы человеческой деятельности, от промышленности и медицины до образования и государственного управления. В то время как многие отрасли увидели потенциал в этой технологии и вложили колоссальные средства во внедрение ИИ-решений в свой бизнес, прогнозы, сделанные с помощью ИИ-алгоритмов, могут быть пугающими и без надлежащей образовательной базы могут привести к недоверию в обществе. Растущая автономия и сложность систем ИИ ставят перед научным сообществом и обществом в целом новые, комплексные вызовы.

Специалисты отмечают, что узкоспециализированный, техноцентричный взгляд на ИИ исчерпал себя. Проблемы, возникающие при внедрении ИИ, зачастую лежат не в технической, а в гуманитарной плоскости. В промышленности чат-боты с искусственным интеллектом оказались некорректными из-за обучающих данных, которые были представлены алгоритму, программное обеспечение для подбора персонала – предвзятым по гендерному признаку, а инструменты оценки рисков, разработанные американским подрядчиком, привели к тюремному заключению невиновных людей [1].

ИИ можно использовать как инструмент для повышения конфиденциальности данных или для выявления угроз, но чаще его рассматривают как угрозу для ИТ-систем [2], например в случаях с биометрической безопасностью и конфиденциальностью. ИИ может стать мишенью для атак с использованием уязвимостей, качественно новых для систем ИИ

(таких как состязательные атаки и атаки с отравлением), а также новым мощным инструментом в руках злоумышленников.

Очевидно, что необходимо тщательно изучать влияние ИИ, следуя глобальным и местным этическим нормам для создания надёжного и ответственного ИИ. Возникающие угрозы не могут быть устранены исключительно силами инженеров и программистов. Именно здесь на передний план выходит концепция человекоориентированного искусственного интеллекта (ЧОИИ), разработка которого немыслима без междисциплинарного подхода.

Для формирования комплексного видения проблемы были изучены и обобщены работы по темам искусственного интеллекта, человеко-машинного взаимодействия, междисциплинарных исследований, этики и философии технологий. Проанализированы работы как отечественных, так и зарубежных авторов. Проведено сопоставление традиционного (техноцентричного) и нового (человекоориентированного) подходов к разработке ИИ для выявления их кардинальных различий по ключевым критериям. На основе анализа литературы выделены и систематизированы основные характеристики ЧОИИ, а также определен вклад различных научных дисциплин в его разработку. Результаты систематизации представлены в виде таблиц. Разработана наглядная схема (рисунок), иллюстрирующая модель междисциплинарного взаимодействия при создании систем ЧОИИ.

ОПРЕДЕЛЕНИЕ И ХАРАКТЕРИСТИКИ ЧЕЛОВЕКООРИЕНТИРОВАННОГО ИИ

В статье [3] «человекоориентированный ИИ» рассматривается как система, основанная на ценностно-ориентированном подходе, то есть при разработке и использовании ИИ в центре внимания находятся человеческие ценности, этические принципы и моральные нормы. Такой подход направлен на создание систем, которые не только эффективны, но и соответствуют интересам и благополучию человека, способствуя его развитию, а не подавляя его.

Коллектив авторов работы [4] рассматривает «человекоориентированный ИИ» через призму ответственного ИИ (responsible AI). Основное внимание уделяется объяснимости ИИ (XAI, от англ. «Explainable Artificial Intelligence» – объяснимый искусственный интеллект). Их определение подразумевает, что ИИ-системы должны быть прозрачными и понятными для людей, способными объяснять свои решения и действия. Также такие системы помогают повысить доверие, обеспечить справедливость и подотчётность, эффективно решать проблемы, которые могут возникнуть в процессе работы ИИ.

Автор статьи [5] анализирует «человекоориентированный ИИ» с точки зрения надежного ИИ (trustworthy AI), основываясь на подходах, принятых в Европейском союзе. Для него основным является обеспечение этических принципов и норм, также он включает в себя такие аспекты, как уважение к человеческой автономии, предотвращение вреда, справедливость и объяснимость. Отмечается, что ИИ должен быть разработан таким образом, чтобы он служил людям, соблюдал их права и не наносил ущерба, являлся надёжным и заслуживающим доверия.

На основе анализа существующих подходов вышеуказанных авторов предлагаем следующее определение: ***человекоориентированный ИИ*** – парадигма проектирования, разработки и внедрения систем искусственного интеллекта, в которой главной целью является не только достижение технических показателей, но и усиление человеческих способностей, уважение человеческого достоинства и прав, обеспечение прозрачности, контролируемости и справедливости принимаемых системой решений в полном соответствии с ценностями и этическими нормами общества.

В отличие от просто «полезного» или «эффективного» ИИ, ЧОИИ подразумевает глубокую интеграцию человеческого фактора на всех этапах жизненного цикла системы. В таблице 1 описаны основные характеристики человекоориентированного ИИ, которые демонстрируют его отличие от традиционных, исключительно технологически ориентированных систем.

Таблица 1

Основные характеристики человекоориентированного ИИ

Характеристика	Описание
Техническая надежность и безопасность	Устойчивость к атакам, сбоям и непредвиденным ситуациям. Способность системы безопасно функционировать в динамичной среде, минимизируя риски для человека и инфраструктуры.
Прозрачность и объяснимость ИИ	Способность системы предоставлять понятные для человека объяснения своих решений и прогнозов. Человек должен понимать, почему ИИ пришел к тому или иному выводу [6].

Продолжение таблицы 1

Характеристика	Описание
Контролируемость и подотчетность	Человек должен сохранять конечный контроль над системой. Должны быть четко определены механизмы ответственности за действия и ошибки ИИ.
Справедливость и непредвзятость	Отсутствие дискриминационных предубеждений (гендерных, расовых, социальных и др.) в данных и алгоритмах. Обеспечение равного и справедливого отношения ко всем группам пользователей [7].
Конфиденциальность и защита данных	Гарантия защиты персональных данных пользователя на всех этапах: сбора, хранения, обработки. Соответствие нормам законодательства (например, общий регламент по защите данных и этическим принципам).
Ориентация на благополучие человека	Цель ИИ – не замена человека, а усиление его когнитивных, творческих и физических способностей. Системы должны способствовать социальному благу и устойчивому развитию.
Уважение человеческой автономии	ИИ не должен манипулировать человеком или ограничивать его свободу выбора. Система должна выступать в роли советчика или ассистента, а не диктатора.
Эволюция и адаптивность	Способность системы к обучению и адаптации на основе взаимодействия с человеком и средой, с учетом обратной связи для постоянного улучшения [8].

Данные характеристики показывают, что человекоориентированный ИИ – это комплексное понятие, выходящее далеко за рамки чисто технических спецификаций. Его реализация требует фундаментального пересмотра самого подхода к разработке ИИ. В основе ЧОИИ лежит не просто создание эффективных алгоритмов, а разработка систем, которые органично и безопасно взаимодействуют с человеком, уважая его ценности и автономию.

Одним из фундаментальных аспектов ЧОИИ является техническая надежность и безопасность. Они формируют базис, без которого все остальные принципы теряют смысл. Система, не способная безопасно функционировать в различных условиях и неустойчивая к сбоям, не может считаться надежным партнёром для человека. На этой основе строятся принципы, связанные с этической и социальной ответственностью [9].

Тесно связаны между собой прозрачность и объяснимость искусственного интеллекта, а также его контролируемость и подотчетность. Объяснимость решений ИИ-системы является необходимым условием для того, чтобы человек мог сохранить контроль над её действиями. Только понимая логику работы системы, пользователь может адекватно её контролировать, вмешиваться в её работу и принимать решения, что, в свою очередь, порождает подотчетность, где ответственность за действия системы четко распределяется между человеком и машиной.

Следующая группа характеристик, таких как справедливость и непредвзятость, а также конфиденциальность и защита данных, относится к этическому измерению ЧОИИ. Использование необъективных данных или алгоритмов может приводить к социально вредным последствиям, в то время как отсутствие надежной защиты данных разрушает основу для долгосрочного и доверительного сотрудничества с пользователем. Соответствие законодательным нормам, таким как GDPR (от англ. General Data Protection Regulation – общий регламент по защите данных, нормативный акт Европейского союза), является обязательным требованием [10].

Наконец, высшей целью ЧОИИ является ориентация на благополучие человека и уважение человеческой автономии. ИИ рассматривается не как замена, а как инструмент для расширения человеческих возможностей. Система должна выступать в роли ассистента или советчика, предоставляя человеку полную свободу выбора и не манипулируя им. Этот подход напрямую связан с эволюцией и адаптивностью ИИ-системы, её способностью постоянно учиться на основе обратной связи от человека, чтобы с каждым новым взаимодействием становиться полезнее и эффективнее, способствуя достижению социального блага.

СРАВНЕНИЕ ТРАДИЦИОННОГО И МЕЖДИСЦИПЛИНАРНОГО ПОДХОДОВ

Междисциплинарный подход – подход к решению научных проблем, основанный на объединении двух и более научных направлений под эгидой какой-либо обобщающей концепции с целью получения новых результатов [11]. Междисциплинарный подход стал научным трендом начиная с последней четверти XX в. и связан со становлением постнеклассической рациональности.

Однако фактически к разработке ИИ до сих пор применяется традиционный или техноцентричный, подход, который фокусируется на оптимизации измеримых метрик: точность, скорость, производительность. Техноцентричный подход привел к впечатляющим технологическим прорывам, но он же и породил проблемы, упомянутые ранее (предвзятость, непрозрачность («черный ящик»), уязвимость и социальное неприятие) [12]. Причина в том, что социальные и этические категории, такие как «справедливость», «доверие» или «благополучие», не могут быть сведены к простой математической функции потерь.

Отличия междисциплинарного подхода от традиционного представлены в таблице 2.

Таблица 2

Сравнение традиционного и междисциплинарного подходов в разработке человекоориентированного ИИ

Критерий	Традиционный (техноцентричный) подход	Междисциплинарный подход
Основная цель	Максимизация технических метрик (точность, скорость)	Усиление человеческих возможностей и социальное благо
Ключевой вопрос	Можем ли мы это сделать?	Должны ли мы это делать и как сделать это правильно?
Роль человека	Пользователь, источник данных, объект воздействия	Партнер, участник проектирования, конечный бенефициар
Процесс разработки	Линейный, ведомый инженерами	Итеративный, с участием социологов, психологов, этиков, юристов
Данные	Рассматриваются как «топливо» для модели	Анализируются на предмет предвзятости, репрезентативности, этичности сбора
Этика и ценности	Рассматриваются как внешнее ограничение или постфактум	Интегрированы в ядро проектирования («Ethics by Design»)
Оценка успеха	Техническая производительность, экономическая выгода	Социальное принятие, доверие пользователей, долгосрочное положительное влияние
Прозрачность	Низкий приоритет, «черный ящик» допустим	Высокий приоритет, объяснимость (XAI) является ключевым требованием

В центре традиционного, техноцентричного подхода лежит стремление к максимизации технических метрик, таких как точность и скорость. Основной вопрос, которым задаются инженеры, звучит так: «Можем ли мы это сделать?». В отличие от него, междисциплинарный подход смещает фокус на усиление человеческих возможностей и социальное благополучие. Здесь основной вопрос кардинально меняется на более глубокий и этический: «Должны ли мы это делать и как сделать это правильно?». Этот сдвиг в целеполагании определяет все последующие этапы разработки.

Изменение роли человека и процесса разработки – различия в целях, которые напрямую влияют на роль человека в процессе. Традиционный подход рассматривает человека, как простого пользователя, источник данных или объект воздействия. Междисциплинарный подход, напротив, делает человека партнёром и активным участником проектирования, а также конечным бенефициаром [13]. Такой сдвиг требует иного процесса разработки, который становится итеративным и включает экспертов из различных областей – социологов, психологов, этиков и юристов, а не только инженеров.

Основным отличием является отношение к данным. В техноцентричном подходе данные воспринимаются как «топливо» для модели, а их этическая сторона редко подвергается сомнению. Междисциплинарный подход анализирует данные на предмет предвзятости, репрезентативности и этичности сбора, чтобы избежать дискриминации. Этика и ценности здесь не являются внешними ограничениями, а интегрированы в ядро проектирования с самого начала, что описывается принципом «этика по замыслу».

Традиционный и междисциплинарный подходы по-разному оценивают успех. Традиционный подход измеряет успех технической производительностью и экономической выгодой. Междисциплинарный подход выходит за рамки этих показателей, оценивая социальное принятие, доверие пользователей и долгосрочное положительное влияние. Высокий приоритет здесь отдаётся прозрачности, и объяснимость ИИ становится основным требованием, чтобы избежать создания «чёрного ящика» и гарантировать, что человек всегда будет понимать и контролировать ИИ.

МОДЕЛЬ МЕЖДИСЦИПЛИНАРНОГО ПОДХОДА

Междисциплинарный подход является этапом эволюции разработки ИИ, на котором технологии служат инструментом для достижения человеческих целей, а не наоборот. Сравнительный анализ подходов доказывает, что для разработки систем в рамках концепции ЧОИИ необходим переход от техноцентризма к междисциплинарному подходу.

Модель, представленная на рисунке 1, демонстрирует, что разработка ЧОИИ является постоянно повторяющимся процессом, основанным на анализе, проектировании, внедрении и контроле и обеспечивающим его постоянное улучшение и соответствие потребностям пользователя. Эта особенность модели также требует формирования междисциплинарного подхода в разработке ЧОИИ.

Междисциплинарный подход предполагает активное сотрудничество специалистов из разных областей на всех этапах создания системы ИИ. Суть подхода заключается в интеграции различных аспектов для решения комплексной задачи. В таблице 3, основанной на приведенной выше модели, представлен жизненный цикл разработки системы человекоориентированного искусственного интеллекта и также показано, что разработка ЧОИИ – не линейный, а циклический и междисциплинарный процесс, состоящий из четырех блоков: проектирование, внедрение, контроль и анализ. Каждый блок имеет свои уникальные этапы и требует участия специалистов из различных областей.

Рис. 1. Модель междисциплинарного подхода в разработке ЧОИИ

Таблица 3

Междисциплинарный подход в жизненном цикле разработки ЧОИИ

Этап жизненного цикла	Участвующие дисциплины	Роль
I. ПРОЕКТИРОВАНИЕ		
1. Формулирование проблемы и цели	Этика и социология	Определяют потенциальное социальное и этическое воздействие.
	Юриспруденция	Анализирует правовые риски и соответствие законодательству.
	Программная инженерия	Технические специалисты определяют техническую реализуемость.
2. Сбор и подготовка данных	Этика и право	Гарантируют законность сбора, анонимность и отсутствие дискриминации.
	Информатика и статистика	Обеспечивают репрезентативность, качество и техническую подготовку данных.
3. Проектирование и разработка модели	Информатика и инженерия	Создают и оптимизируют алгоритмы.
	Дизайн и психология	Разрабатывают пользовательский интерфейс и логику взаимодействия, учитывая человеческие ценности.
	Этика	Внедряет ценностные и этические ограничения в архитектуру модели.
II. ВНЕДРЕНИЕ		
4. Формирование регламентирующей документации	Юриспруденция	Разрабатывают внутренние правила и инструкции.
	Менеджмент и маркетинг	Обеспечивают правильную коммуникацию и внедрение.
5. Доведение документации до сотрудников	Менеджмент и маркетинг	Организуют обучение и распространяют информацию.
	Психология	Обеспечивает усвоение и принятие новых правил.
III. КОНТРОЛЬ		
6. Контроллинг бизнес-процессов на основе данных ИТ-систем	Информатика и инженерия	Создают механизмы мониторинга.
	Анализ данных	Анализирует эффективность и производительность.
7. Контроль показателей	Статистика	Измеряют ключевые метрики.
	Менеджмент	Принимают решения на основе данных.
IV. АНАЛИЗ		
8. Анализ несоответствий и их последствий	Социология и этика	Выявляют отклонения и потенциальные предвзятости.
	Анализ данных	Оценивает социальные и этические последствия.
9. Анализ показателей	Статистика	Сравнивает фактические данные с целевыми.
	Анализ данных	Оценивает эффективность решения.

Процесс разработки ЧОИИ можно представить следующим образом.

На первом этапе ***«Проектирование»*** закладывается вся основа будущего проекта. Менеджмент и специалисты по анализу данных совместно работают над формализацией стратегии, определяя общие цели и место системы ИИ в компании. Затем они, вместе с инженерами-разработчиками ИИ, переходят к проектированию архитектуры и логики алгоритмов. На этом этапе также подключаются HR-специалисты и менеджмент для проектирования организационной структуры, что необходимо для подготовки компании к внедрению новой технологии.

Ключевую роль играют информатика, математика и статистика, которые проводят имитационное моделирование и функционально-стоимостный анализ, прогнозируя эффективность будущей модели. Этап завершается разработкой технического задания, при этом программные инженеры и юристы учитывают все требования, включая правовые нормы, например, в части защиты данных.

Следующий блок ***«Внедрение»*** сосредоточен на подготовке и запуске системы ИИ. На этом этапе юриспруденция, менеджмент и отдел по связям с общественностью работают над формированием регламентирующей документации, связанной с использованием и обслуживанием модели. Затем менеджмент, HR и специалисты по коммуникациям обеспечивают доведение этой документации до сотрудников, организуя обучение и тренинги по взаимодействию с ИИ. Психология на этом этапе помогает обеспечить усвоение и принятие новых правил персоналом, а также снизить сопротивление изменениям.

Блок ***«Контроль»*** отвечает за мониторинг и оценку работы системы. Информатика и инженерия создают механизмы для контроллинга производительности модели, а специалисты по анализу данных анализируют её эффективность. На этапе контроля показателей статистика измеряет ключевые метрики (например, точность или время отклика), а менеджмент принимает на их основе решения о доработке или масштабировании системы. Этот блок является важным для сбора фактических данных о работе модели.

И, наконец, блок ***«Анализ»*** замыкает цикл. Социология и этика, совместно с анализом данных, анализируют несоответствия и их последствия, выявляя отклонения и потенциальные предвзятости в поведении модели, а также оценивая их социальное и этическое влияние. На этапе анализа показателей статистика и анализ данных сравнивают фактические результаты с целевыми, оценивая общую эффективность решения. Полученные в ходе анализа данные и выводы затем возвращаются на этап «Проектирования» для итеративной доработки модели, и цикл повторяется.

Таким образом, междисциплинарный подход в разработке ЧОИИ основан на непрерывном процессе, где каждый этап тесно связан с предыдущим и последующим, обеспечивая постоянное совершенствование и адаптацию системы к изменяющимся условиям и требованиям.

Результаты данного исследования убедительно демонстрируют, что переход от техноцентричной парадигмы к концепции человекоориентированного искусственного интеллекта является не просто академической дискуссией, а насущной необходимостью, продиктованной как технологическими, так и социальными вызовами. Традиционный подход, ориентированный исключительно на достижение максимальных технических метрик, таких как точность, скорость и производительность, привел к созданию мощных, но в то же время потенциально опасных и социально безответственных систем.

Недостатком техноцентризма является то, что он сводит сложную человеческую реальность к набору измеряемых данных, упуская из виду такие не поддающиеся количественной оценке категории, как справедливость, достоинство, доверие и автономия. В результате алгоритмы, обученные на предвзятых данных, не просто воспроизводят, но и усиливают существующие социальные предубеждения, создавая эффект «алгоритмической предвзятости». Такой подход формирует «черный ящик», который способен принимать жизненно важные решения (например, в сфере правосудия или медицины), но не может объяснить свою логику, подрывая доверие общества к технологиям. Создается так называемый «этический долг» ИИ, аналогичный «техническому долгу» в программировании, где нерешенные проблемы на ранних этапах приводят к огромным издержкам в будущем.

Именно здесь на первый план выходит междисциплинарный подход. Взаимосвязь между дисциплинами на каждом этапе жизненного цикла, демонстрирует, что успех ЧОИИ возможен только при непрерывном сотрудничестве. На этапе анализа этики и социологи работают вместе с инженерами, чтобы не только определить техническую задачу, но и оценить её потенциальное социальное и этическое воздействие. Психологи и специалисты по дизайну пользовательского опыта помогают создавать интерфейсы, которые уважают человеческую автономию (а не манипулируют ею), обеспечивая ин-

туитивную понятность и контроль. Юристы, в свою очередь, гарантируют, что система соответствует законодательным нормам, таким как «Общий регламент по защите данных», защищая права и конфиденциальность пользователей.

Однако, внедрение парадигмы ЧОИИ сопряжено с определенными вызовами. Одним из наиболее острых является проблема перевода абстрактных гуманитарных понятий в конкретные технические требования. Например, как измерить и закодировать «справедливость» или «доверие» в математическую модель? Требуется разработка новых метрик и инструментов оценки, которые будут учитывать не только техническую производительность, но и социальное принятие. Кроме того, для успешного междисциплинарного сотрудничества необходимы новые образовательные программы и организационные структуры, которые будут способствовать открытому диалогу и взаимопониманию между специалистами из разных областей.

Тем не менее, преимущества междисциплинарного подхода значительно перевешивают его сложности. ЧОИИ, разработанный в рамках междисциплинарной парадигмы, становится более надежным, справедливым и прозрачным. Он не только снижает риски возникновения социальных и этических проблем, но и открывает новые возможности для создания технологий, которые по-настоящему служат на благо человечества, усиливая наши способности, а не подменяя их.

ЗАКЛЮЧЕНИЕ

В результате проведенной концептуализации человекоориентированного искусственного интеллекта обоснована критическая важность междисциплинарного подхода для его успешной разработки, подтверждена необходимость перехода к ЧОИИ от традиционного техноцентричного подхода, игнорирующего социальные, этические и правовые аспекты и неизбежно приводящего к созданию систем, которые могут быть предвзятыми, небезопасными и неприемлемыми для общества.

Опыт последних лет наглядно демонстрирует, что техноцентричный ИИ, который не учитывает человеческий фактор, не может считаться надежным и ответственным инструментом. Напротив, ЧОИИ как система, которая не только эффективно выполняет свои технические функции, но и способствует усилению человеческих возможностей, уважает автономию человека и обеспечивает прозрачность и подотчетность, является единственным жизнеспособным путем для развития искусственного интеллекта.

Проведенный сравнительный анализ выявил кардинальные отличия между традиционным и междисциплинарным подходом, показав сдвиг в фокусе с оптимизации метрик к созданию социальной ценности через интеграцию знаний и экспертизы из инженерии, информатики, психологии, этики, социологии, права и дизайна на всех этапах жизненного цикла продукта.

Таким образом, междисциплинарный подход должен стать не просто рекомендованной практикой, а обязательным стандартом в разработке ИИ. Только в условиях такого сотрудничества возможно создание систем, которые будут действительно служить на благо человека, соответствовать его ценностям и способствовать построению более справедливого, безопасного и технологически развитого общества.

Дальнейшие исследования должны быть направлены на разработку практических методологий и инструментов для реализации междисциплинарного подхода в академической и индустриальной среде.

СПИСОК ЛИТЕРАТУРЫ

1. Dressel J., Farid H. The accuracy, fairness, and limits of predicting recidivism // Sci. Adv. 2018. Vol. 4. Art. eaao5580. DOI: 10.1126/sciadv.aao5580.
2. Berghoff C., Neu M., von Twickel A. Vulnerabilities of connectionist AI applications: evaluation and defense // Front. Big Data. 2020. Vol. 3. Art. 213005576. DOI: 10.3389/fdata.2020.00023.
3. Мамина Р.И., Ильина А.В. Новации современного искусственного интеллекта: ценностно-ориентированный подход // ДИСКУРС. 2025. Т. 11. № 1. С. 16–30. DOI: 10.32603/2412-8562-2025-11-1-16-30.
4. Arrieta A.B., Díaz-Rodríguez N., Del Ser J., Bennetot A., Tabik S., Barbado A., et al. Explainable Artificial Intelligence (XAI): concepts, taxonomies, opportunities and challenges toward responsible AI // Inf. Fusion. 2020. Vol. 58. P. 82–115. DOI: 10.1016/j.inffus.2019.12.012.
5. Smuha N. A. The EU approach to ethics guidelines for trustworthy artificial intelligence // CRi-Comput. Law Rev. Int. 2019. Vol. 20. P. 2194–4164. DOI: 10.9785/cri-2019-200402.
6. Куликов В.П., Куликова В.П., Кухаренко Е.В. Интеграция ИИ и СППР в междисциплинарной подготовке

ИТ-специалистов: кейс-подход к разработке интерфейсов прямого манипулирования // Вестник Северо-Казахстанского Университета им. М. Козыбаева. 2025. Т. 66. № 2. С. 207–219. DOI: 10.54596/2958-0048-2025-2-207-219.

7. Вальченко Ю., Кашевник А.М. Современные подходы к построению контекстноориентированных систем в интеллектуальных пространствах // Тр. СПИИРАН. 2011. Вып. 4 (19). С. 102–127. DOI: https://doi.org/10.15622/sp.19.6

8. Карпов А.А. Эволюция развития человеко-машинных интерфейсов и современные технологии искусственного интеллекта // XXVII Годичная научная международная конференция Института истории естествознания и техники им. С. И. Вавилова РАН, Москва- Санкт-Петербург, 17–21 мая 2021 года. – Москва: Институт истории естествознания и техники им. С. И. Вавилова РАН, 2021. С. 643–646.

9. Лори Н.Ф. Междисциплинарность в инженерном образовании: тенденции и концепции // Инженерное образование. 2014. № 14. С. 30–37.

10. Dixon W., Eagan N. 3 ways AI will change the nature of cyber attacks // World Economic Forum. Davos, Switzerland, 2019. URL: https://www.weforum.org/stories/2019/06/ai-is-powering-a-new-generation-of-cyberattack-its-also-our-best-defence/ (дата обращения: 22.01.2025).

11. Gauthier J., Levy R. Linking artificial and human neural representations of language // arXiv preprint arXiv:1910.01244. 2019. P. 1–15.

12. Jiang R., Al-maadeed S., Bouridane A., Crookes D., Beghdadi A. (Editors). Biometric security and privacy: opportunities and challenges in the big data era // Cham, Switzerland: Springer International Publishing, 2017. 423 p.

13. Qiu S., Liu Q., Zhou S., Wu C. Review of artificial intelligence adversarial attack and defense technologies // Appl. Sci. 2019. Vol. 9. Art. 909. DOI: 10.3390/app9050909.

УДК: 130.2, 004

Эволюция культуры информационной безопасности: организационно-технические и социокультурные факторы

P.G. Bylevskiy

The Evolution of Information Security Culture: Organizational, Technical, and Socio-Cultural Factors

Abstract. *The article examines the evolution of information security culture under the influence of changing organizational, technical and socio-cultural factors in the development of computer network technologies. It analyzes the relationship between the development of equipment, software, network solutions, Internet communications and digital services, the complication of the nature of use, the increase in the number and diversity of users, protected/attacked values (quantitatively, qualitatively and by significance), threats, risks and means of protection. It reveals an increase in the importance of socio-cultural factors, reaching a maximum as a result of digital transformation. The article studies the development of narrowly professional activities to protect corporate secrets from technical intelligence to the general civil culture of information security of everyday life, traditional values and cultural identity.*

Keywords: *culture of information security, evolution of computer network technology, Internet communications, modern digital technologies, values, threats.*

П.Г. Былевский
Кандидат философских наук, доцент ВАК 2.3.6. «Методы и системы защиты информации, информационная безопасность», доцент кафедры международной информационной безопасности Московского государственного лингвистического университета, старший преподаватель Российского государственного социального университета
E-mail: pr-911@yandex.ru

Аннотация. *В статье исследуется эволюция культуры информационной безопасности под влиянием изменяющихся во времени организационно-технических и социально-культурных факторов развития компьютерно-сетевых технологий. Проанализирована взаимосвязь развития оборудования, программного обеспечения, сетевых решений, интернет-коммуникаций и цифровых сервисов, усложнения характера пользования, увеличения количества и разнообразия пользователей, защищаемых/атакуемых ценностей (количественно, качественно и по значимости), угроз, рисков и средств защиты. Выявляется увеличение значения социально-культурных факторов, достигающего максимума в результате цифровой трансформации. Изучается развитие узкопрофессиональной деятельности по защите корпоративной тайны от средств технической разведки до общегражданской культуры информационной безопасности повседневного быта, традиционных ценностей и культурной идентичности.*

Ключевые слова: *культура информационной безопасности, эволюция компьютерно-сетевых технологий, интернет-коммуникации, современные цифровые технологии, ценности, угрозы.*

ВВЕДЕНИЕ

Задачей настоящей статьи является исследование эволюции культуры информационной безопасности на основе формирования организационно-технических и социально-культурных факторов развития компьютерно-сетевых технологий. Безопасность современных цифровых технологий представляет собой не только профессиональную деятельность, но и общегражданскую культуру, что констатируется «Концепцией формирования и развития культуры информационной безопасности граждан Российской Федерации» Правительства РФ: «Культура информационной безопасности – это совокупность сформированных знаний, умений и навыков по вопросам информационной безопасности, обеспечивающая безопасное пребывание гражданина Российской Федерации в информационном пространстве»[1].

За исходную позицию развития культуры информационной безопасности можно принять узкопрофессиональную деятельность защиты конфиденциальных сведений от технических средств разведки [1]. Развитие компьютерно-сетевых технологий и их применений расширило информационную безопасность технологий до общегражданской культуры повседневного быта, защиты традиционных ценностей и культурной идентичности.

[1] Концепция формирования и развития культуры информационной безопасности граждан Российской Федерации (утверждена распоряжением Правительства Российской Федерации от 22 декабря 2022 г. № 4088-р (Дата опубликования: 23.12.2022). С.2. URL: http://publication.pravo.gov.ru/Document/View/0001202212230035?index=3 (Дата обращения 14.07.2025).

ОСНОВАНИЯ ДЛЯ ОПРЕДЕЛЕНИЯ ЭТАПОВ ЭВОЛЮЦИИ ИНФОРМАЦИОННОЙ БЕЗОПАСНОСТИ

Термин «информационная безопасность» входит в массовое употребление в 1990-е годы в связи с началом гражданского распространения персональных настольных компьютеров и, на этой основе, распространением глобальной массовой публичной сети интернет [2]. «Конверсия» компьютерных технологий для массовых гражданских применений на первых порах требовала в основном трансляции новым пользователям профильных технических компетенций, без учёта особенностей обеспечения конфиденциальности, свойственной сотрудникам спецслужб и военным.

Весьма показательным можно считать опыт Российского государственного гуманитарного университета (РГГУ), где в 1990 году для подготовки гражданских специалистов по организации и технологиям защиты информации был создан факультет защиты информации, в 2001 году реорганизованный в Институт информационных наук и технологий безопасности. Впервые были разработаны профильные учебные дисциплины в «гражданских» областях защиты информации в автоматизированных системах обработки данных и управления, а также криптографии (В.А. Герасименко, А.Н. Лебедев, А.Ю. Щербаков и др.).

Базовым, организационно-техническим фактором превращения информационной безопасности из узкопрофессиональной деятельности в общегражданскую культуру было развитие компьютерно-сетевых технологий и их применений. Эволюция компьютерно-сетевых технологий проходила несколько этапов, характеризовавшихся рядом существенных параметров: «поколений» оборудования и сетевых решений, интернет-коммуникаций. Соответственно происходила эволюция информационной безопасности [3], включая такие социально-культурные факторы, как количество и состав пользователей, характер пользования, обрабатываемые (атакуемые / защищаемые) ценности, угрозы, риски и средства защиты.

Основаниями для определения этапов и хронологических границ эволюции культуры информационной безопасности служат ключевые **организационно-технические и социокультурные факторы и аспекты** в России (до 1991 года в СССР). Значимыми **факторами культуры информационной безопасности**, определяющими её эволюцию, выступают развитие компьютерно-сетевых технологий и интернет-коммуникаций, нормативно-правовые основы, защищаемые/атакуемые ценности, количество и состав пользователей, характер пользования, угрозы и риски, типы нарушителей, средства атак и защиты.

К существенным **аспектам культуры информационной безопасности** относятся обучение и воспитание, психологические качества (осторожность, эмоциональная устойчивость, дисциплинированность и др.), усвоение этических норм, мировоззрение, культурная идентичность, традиционные ценности, профессиональный/общегражданский характер компетенций. Увеличение доли и значения социокультурных факторов, и, соответственно, аспектов информационной безопасности обусловливает её развитие как культуры, от профессиональной до общегражданской.

Эволюция информационной безопасности происходила на основе развития, усложнения, расширения и интенсификации применения организационно-технических факторов — компьютерно-сетевых технологий, полного охвата общества и всех граждан, всех отраслей и быта, одновременно наполняясь социально-культурным содержанием. При увеличении на порядки объёмов и сложности высокотехнологичного, компьютерно-сетевого оборудования в процессе этой эволюции удельный вес и относительное значение организационно-технических факторов, напротив, сокращались относительно социокультурных.

Новые гражданские применения компьютерно-сетевых технологий актуализировали акцент на различных аспектах классической «триады» информационной безопасности: государственные и корпоративные компьютерные системы — на конфиденциальности, юридически значимый документооборот — на целостности (электронная подпись), публичные массовые цифровые сервисы — на доступности, бесперебойности работы.

Прослеживаются следующие последовательные изменения взаимосвязанных **организационно-технологических и социально-культурных факторов информационной безопасности**:

- организационно-технологического базиса (оборудования, программного обеспечения, типов сетевых соединений и доступа, архитектуры решений);
- «поколений» интернет-коммуникаций (web 1.0, 2.0, 3.0) и сервисов (от электронных к цифровым) [4];
- количества и разнообразия пользователей (чья природа социально-культурна), самого характера пользования;
- защищаемых/атакуемых ценностей (количественно, качественно и по значимости);

• угроз (типов нарушителей и др.);
• рисков потенциального ущерба.

ОРГАНИЗАЦИОННО-ТЕХНОЛОГИЧЕСКИЕ ФАКТОРЫ ИНФОРМАЦИОННОЙ БЕЗОПАСНОСТИ

Базисом, главным организационно-технологическим фактором эволюции культуры информационной безопасности является развитие, смена «поколений» электронно-вычислительной техники — компьютерного оборудования, сетевых решений. Выделяются следующие **виды компьютерного оборудования**, создание и массовое применение которых приводило к качественным переменам в характере пользования и задачах информационной безопасности:

• индустриальные централизованные ЭВМ, «мейнфреймы» (государственные, ведомственные, корпоративные) [5][2];
• настольные персональные компьютеры [6];
• мобильные персональные компьютеры (смартфоны, планшеты) [7];
• «интернет вещей», специализированные компьютерные устройства (промышленные и бытовые, стационарные и носимые).

Новые виды компьютерно-сетевых решений не заменяют, не вытесняют старые, а создают новые «геологические слои», в то же время вызывая существенные перемены на нижних уровнях, в технологическом базисе. Важнейшим направлением развития компьютерного оборудования стали сетевые решения (возможности передачи компьютерных данных), организационно-технологические условия совместной (одновременно или поочередно) работы пользователей на одном или нескольких компьютерах.

Поочерёдная или одновременная работа нескольких пользователей, использование записей данных на перфокартах, в том числе на других компьютерах, постепенно развиваются в масштабные, многофункциональные сетевые решения, сложно структурированные и охватывающие обширные территории (вплоть до глобального уровня).

Развитие сетевых решений происходило в следующей последовательности:

• передача записей на носителях (перфокартах, перфолентах, магнитных, электромагнитных накопителях и т.д.), позже посредством специальной связи по защищённым каналам;
• проводные локальные компьютерные сети разного уровня (корпоративные, территориальные и др.);
• низкоскоростная аналоговая проводная телефонная и беспроводная радиосвязь;
• широкополосная (быстрая) передача данных по проводной (оптоволоконной и др.) и беспроводной связи (мобильной, спутниковой и др.).

Ниже приведем **усложняющиеся виды компьютерно-сетевых сервисов**:

• для служебного пользования (государственных, оборонных нужд, научных исследований, разработок и т.п.);
• гражданские некоммерческие локально-сетевые профессиональные и любительские (модемная связь по аналоговой телефонии) применения: делопроизводство, документооборот, редакционно-издательское дело, текстовое общение и т.п. [8];
• доступ к публичным интернет-сервисам web 1.0 («трансляция» на сайтах для просмотра и сохранения преимущественно негосударственного, некоммерческого контента; поисковые машины, электронная почта, чаты, тематические форумы обсуждений, одноуровневые файлообменные сети);
• быстрый рост объёмов и разнообразия индивидуального, корпоративного и государственного социально-культурного контента на интернет-ресурсах (книг, прессы, изображений, звукозаписей, видео, игр и т.п. [9]), преимущественно некоммерческого;
• формально бесплатные для пользователей интерактивные массовые цифровые интернет-сервисы web 2.0: социальные сети для публичного мультимедийного общения (отложенного и в реальном времени) и творчества (платформы блогов: текстов, изображений, видео и трансляций);
• мобильные интернет-сервисы (приложения для планшетов, смартфонов и гаджетов);
• автоматизированный мониторинг, аналитика и управление «интернетом вещей» web 3.0 (потенциально всей социальной, индустриальной и бытовой техносферой).

СОЦИАЛЬНО-КУЛЬТУРНЫЕ ФАКТОРЫ ИНФОРМАЦИОННОЙ БЕЗОПАСНОСТИ

Разработка компьютерно-сетевых (как, впрочем, и любых других) технологий тесно связана с организационным фактором, непосредственным и после-

[2] Вопреки мнимой очевидности, массовое распространение персональных настольных, мобильных компьютеров и «интернета вещей» не отрицает и не преуменьшает, но напротив, усиливает роль централизованных индустриальных компьютерных систем («суперкомпьютеров» и центров обработки данных) как организационно-технического ядра, регулирующего публичные интернет-сервисы.

довательным взаимодействием создателей и пользователей. Социально-культурная природа людей, возможности абстрагирования от которой ограничены, отражается в изменении сервисов, характера пользования компьютерно-сетевыми технологиями. Компьютерные сервисы по мере увеличения количества пользователей до массовых масштабов, увеличения времени и вовлечённости пользования оказываются сопряжёнными с многими социокультурными ценностями вплоть да самого высокого уровня [10].

Наряду с увеличением перечня, охвата и возможностей **компьютерно-сетевых интернет-сервисов** следует отдельно указать возможности их **коммерциализации** (легальной либо нелегальной):

• коммерческие публичные массовые интернет-сервисы: банкинг, торговля, реклама и маркетинг;

• скрытое использование крупнейшими корпорациями возможностей интернета в своих коммерческих, политических и других интересах: торговля большими данными, манипулирование массовым сознанием, вмешательство во внутренние дела других стран [11] и т.п.

Преимущественно социокультурными являются такие факторы культуры информационной безопасностей, как пользователи и характер пользования компьютерно-сетевыми технологиями и интернет-коммуникациями, решающим образом связанные с типами личностей, мировоззрением, ценностями, интересами, потребностями, социальным положением и отношениями с другими людьми.

После принятия «Концепции формирования и развития культуры информационной безопасности граждан Российской Федерации» Правительства РФ[3] профильные дисциплины были разработаны и внедрены во многих высших учебных заведениях, учитывая предшествующий опыт преподавания различных аспектов этого предмета в образовательных организациях разного уровня.

Рассмотрим **эволюцию количества и состава пользователей, характера пользования**:

• немногочисленные профессионалы высшей квалификации, обладающие специальным допуском к конфиденциальным сведениям, включая корпоративную, военную тайну, корпоративные секреты;

• минимальное количество гражданских пользователей выстраиваемых интернет-коммуникаций, «любителей», преимущественно специалистов по технике и компьютерным технологиям;

• массовое пользование для профессиональной и бытовой связи (телефонной мобильной, электронной почты, мессенджеров и т.п.), ознакомления с публикациями электронных документов на интернет-ресурсах (текстами, изображениями, звукозаписями и трансляциями);

• вовлечение большинства граждан в постоянное многочасовое пользование интернет-коммуникациями, становящееся одной из первоочередных социокультурных потребностей, вплоть до избыточности (наносящей вред);

• превращение почти всех граждан не только в пользователей, но и в постоянный объект автоматизированной сетевой генерации и анализа данных, обратного управления цифровой социальной техносферой, сообществами и личностью [12].

Следующим важным социокультурным фактором информационной безопасности являются ценности, напрямую связанные с использованием компьютерно-сетевых технологий и интернет-коммуникаций. По мере увеличения количества пользователей сетевых компьютерных сервисов и видов пользования увеличиваются объёмы, разнообразие и значимость ценностей (прямо и косвенно создаваемых, обрабатываемых и потребляемых), которые потенциально могут быть отчуждены. Соответственно возникают и развиваются угрозы и риски, и, как встречная необходимость, потребность в обеспечении информационной безопасности, вначале как узкой организационно-технической профессиональной специализации, а постепенно и как массовой общегражданской культуры.

Прослеживается расширение **перечня атакуемых / защищаемых ценностей**, создаваемых, обрабатываемых и используемых посредством компьютерно-сетевых технологий (интернет-коммуникаций), потенциальных мишеней для злоумышленников (нарушителей) и объектов защиты легальными пользователями:

• корпоративная тайна, оборудование и каналы специальной связи;

• программное обеспечение и электронные документы;

• конфиденциальная информация, репутации;

• денежные средства, объекты авторского права, другое имущество;

• поведение, потребности, убеждения, традиции, идентичность, мировоззрение личности и общностей;

• политическая стабильность, общественная безопасность, государственный суверенитет и без-

[3] Концепция формирования и развития культуры информационной безопасности граждан Российской Федерации (утверждена распоряжением Правительства Российской Федерации от 22 декабря 2022 г. № 4088-р (Дата опубликования: 23.12.2022). С.2. URL: http://publication.pravo.gov.ru/Document/View/0001202212230035?index=3 (Дата обращения 14.07.2025).

опасность, право на жизнь, личное и семейное благополучие.

Расширение перечня и увеличение значимости ценностей, которые могут быть атакованы посредством компьютерно-сетевых технологий (интернет-коммуникаций), определяет эволюцию рисков и угроз информационной безопасности. Разрастается перечень, увеличиваются размеры, возрастает вероятность нанесения разнообразного по видам и значительности **ущерба**, такого как:

• разглашение военных и т.п. тайн (позже также и коммерческих), связанное с высокотехнологичными средствами разведки;

• временная неработоспособность, несанкционированное изменение, порча и уничтожение компьютерного оборудования и программного обеспечения, электронных документов;

• несанкционированный доступ и использование конфиденциальных сведений, нарушение авторского права;

• дистанционные (посредством компьютерно-сетевых технологий и интернет-коммуникаций) кражи денежных средств, незаконное присвоение и порча имущества;

• манипуляции сознанием (в том числе внешние), вовлечение в деструктивную и экстремистскую деятельность, дестабилизация общества, ослабление и разрушение государства;

• девальвация и подмена традиционных ценностей, социокультурной идентичности, фальсификация истории;

• торможение развития российских высокотехнологичных отраслей из-за антироссийских санкций недружественных стран;

• ослабление государства, утрата суверенитета и территориальной целостности, массовые бедствия и гибель граждан.

Соответственно развитию характера пользования и перечня ценностей, связанных с применением компьютерно-сетевых технологий, интернет-коммуникаций (являющихся трансграничными), в контексте динамики противоречий и конфликтов в международных отношениях наблюдается увеличение перечня **субъектов угроз (нарушителей информационной безопасности, злоумышленников)** и повышение уровня возможностей, согласно официальной классификации[4]:

• шпионы (сотрудники спецслужб других стран, корпоративных разведок конкурирующих компаний, ведущие разведку с помощью высокотехнологичных средств);

• хакеры-любители (действующие ради самовыражения и из хулиганских побуждений);

• массовая негосударственная интернет-преступность (криминальный бизнес);

• «телефонные» и интернет-мошенники, «социальные инженеры», деструктивные сообщества, политические экстремисты, террористы;

• компании и некоммерческие организации, осуществляющие цензуру и «мягкое» вмешательство во внутренние дела других стран;

• недружественные страны и организации, занимающиеся дезинформацией, деструктивными воздействиями на традиционные ценности российских граждан, побуждением их к совершению диверсий и террористических актов [13], а также лица и организации, признанные в Российской Федерации иностранными агентами.

ЭТАПЫ ЭВОЛЮЦИИ ИНФОРМАЦИОННОЙ БЕЗОПАСНОСТИ

Смена «поколений» электронно-вычислительной техники, компьютерно-сетевых технологий, интернет-коммуникаций и сервисов служит базовым организационно-технологическим фактором, с которым взаимосвязаны такие важные в социокультурном плане факторы, как характеристики пользователей и характер пользования, атакуемые ценности, средства атаки и зашиты, выступает базисом для определения хронологических границ эволюции культуры информационной безопасности (рубежные даты округлены по десятилетиям):

• 1-й этап: «досетевой», закрытый непубличный период централизованных суперкомпьютеров (1940 – 1980-е гг.) [14];

• 2-й этап: создание и расширение многоуровневых публичных сетей (интернет-коммуникаций) благодаря массовому производству настольных персональных компьютеров (1990 – 2000 гг.);

• 3-й этап: глобальные «трансляционные» интернет-коммуникации, технологии web 1.0. (2000 – 2010 гг.);

• 4-й этап: интерактивные интернет-коммуникации, цифровые сервисы web 2.0 (2010 – 2014[5] гг.);

• 5-й этап: интернет-коммуникации web 3.0 социальной «цифровой техносферы»; большинство

[4] Уровни возможностей нарушителей / ФАУ «ГНИИИ ПТЗИ ФСТЭК России». 14.07.2025. URL: https://bdu.fstec.ru/threat-section/potential (Дата обращения 14.07.2025).

[5] Эта рубежная дата нового этапа эволюции культуры информационной безопасности определена в связи с началом развёртывания в сфере высоких технологий и публичных цифровых сервисов антироссийских санкций недружественными странами и организациями в связи с вхождением Крыма в состав России.

граждан является не только пользователями, но и объектами генерации данных, а также обратных воздействий вне зависимости от их воли и осведомлённости (с 2014 г. по настоящее время).

Учитывая организационно-техническое развитие компьютерно-сетевых технологий, их разработок, применений и пользования, сопряжённых угроз, рисков и средств защиты, исторический анализ позволяет выделить **пять основных этапов развития главных факторов культуры информационной безопасности**. Каждый из этапов комплексно характеризуется специфическими сочетаниями электронно-вычислительных технологий (оборудования, программного обеспечения, архитектуры решений, способом сетевого доступа и т.п.), характером и масштабами их применений и пользования, видами и значимостью защищаемых/атакуемых отчуждаемых ценностей, параметрами возможного ущерба, типами нарушителей, их инструментария и средств обеспечения информационной безопасности (см. Таблицу 1).

Таблица 1

Этапы эволюции организационно-технических и социально-культурных факторов информационной безопасности

	Факторы	Хронологические границы этапов эволюции информационной безопасности				
		1940-1980 гг.	1990-2000 гг.	2000-2010 гг.	2010-2014 гг.	С 2014 г.
I.	**Организационно-технологические факторы**					
1.	**Компьютерное оборудование**	Крупные централизованные ведомственные, корпоративные ЭВМ	Настольные персональные компьютеры	Мобильные персональные компьютеры и центры обработки данных («облачные» сервисы)	Промышленный и бытовой «интернет вещей», «большие данные», распределённые вычисления (в т.ч. блокчейн)	Формирование целостной цифровой инфраструктуры как социальной среды
2.	**Сетевые соединения, типы доступа**	Проводная защищённая, спецсвязь	Локальные сети; создание и расширение публичных интернет-коммуникаций с низкоскоростным аналоговым доступом	Массовое распространение быстрого широкополосного проводного доступа в интернет (web 1.0)	Повсеместный постоянный широкополосный беспроводной доступ пользователей в глобальную сеть интернет (web 2.0)	Повсеместный постоянный широкополосный беспроводной доступ к цифровой инфраструктуре (web 3.0)
3.	**Пользовательские интернет-сервисы**	Не публичные, конфиденциальные, для служебного пользования	«Трансляционные» ресурсы, минимум владельцев, контента, пользователей	Коммерциализация интернета: банкинг, торговля, реклама, дистанционная работа и обучение, культурный контент	Интерактивный публичный интернет (социальные сети, видеохостинги, блоги, массовые цифровые сервисы)	Глобализация централизуемой автоматизации социальной техносферы («искусственный интеллект»)
II.	**Социально-культурные факторы**					
4.	**Пользователи, характер пользования**	Немногочисленные высокие профессионалы, специальный допуск	«Любители», преимущественно специалисты по технике и компьютерным технологиям	Массовое использование телекоммуникаций (мобильная связь, электронная почта, мессенджеры)	Превращение большинства граждан в постоянных пользователей, а пользования – в одну из первоочередных социально-культурных потребностей	Превращение большинства граждан в постоянный объект автоматизированных генерации и анализа данных, управления цифровой техносферой

Продолжение таблицы 1

	Факторы	Хронологические границы этапов эволюции информационной безопасности				
		1940-1980 гг.	1990-2000 гг.	2000-2010 гг.	2010-2014 гг.	С 2014 г.
5.	**Атакуемые / защищаемые ценности**	Корпоративная и коммерческая тайна	Программное обеспечение и электронные документы	Деньги, репутации, конфиденциальная информация	Гражданские права, личность, деньги, имущество, политическая стабильность, общественная безопасность	Государственный суверенитет и безопасность, права на жизнь и благополучие, традиционные ценности, социально-культурная идентичность
6.	**Субъекты угроз (нарушители)**	Спецслужбы других стран и конкуренты	Хакеры-«любители»	Массовая негосударственная интернет-преступность	Телефонные и интернет-мошенники, «социальные инженеры», деструктивные сообщества, политические экстремисты, организаторы вмешательств во внутренние дела, дезинформаторы из недружественных стран и организаций	Недружественные страны и экстремистские организации, организаторы деструктивных воздействий на традиционные ценности, российских граждан, «иноагенты»
7.	**Риски потенциального ущерба**	Разглашение корпоративной тайны, кража коммерческих ноу-хау	Неработоспособность оборудования, порча программного обеспечения и документов	Дистанционные кражи денежных средств, хищение конфиденциальных сведений	Хищения денег и незаконное присвоение имущества, вовлечение в деструктивную и экстремистскую деятельность, общественная дестабилизация, ослабление и разрушение государства	Разрушение и распад государства, утрата суверенитета и территориальной целостности, массовая гибель и бедствия граждан

ЗАКЛЮЧЕНИЕ

Создание новых видов, «поколений» электронно-вычислительной техники и сетевых решений как организационно-технического базиса информационной безопасности связано с социально-культурными факторами (количество и состав пользователей, характер пользования, атакуемые и защищаемые ценности, угрозы, риски и средства защиты).

Взаимодействие организационно-технических и социально-культурных факторов информационной безопасности достигает высокого уровня в цифровой трансформации, при преобразовании индустриальной и бытовой социальной среды в целостную компьютеризованную цифровую инфраструктуру (технологически символизируемую триадой «интернет вещей» — «большие данные» — «искусственный интеллект») [15]. Кроме того, цифровая трансформация предоставляет организаторам сетевых сервисов потенциальные возможности автоматизации максимального охвата данных (посредством анализа «больших данных») и управления подключаемыми компьютерными устройствами в режиме реального времени.

Таким образом, в современных технологиях начал проявляться «роевой» принцип: функционально как на едином компьютере, вычисления

осуществляются на многих миллионах единиц сетевого компьютерного оборудования. Несмотря на мнимую автономность подключенных к сетям разных уровней компьютерных устройств, посредством дистанционных центров обработки данных на основе отдельных компьютерных устройств складывается многоуровневая сетевая система с централизованным сбором данных и обратным управлением.

Одним из наиболее важных следствий рассмотренных процессов цифровой трансформации является образование единой глобальной цифровой социальной техносферы и превращение граждан не только в постоянных пользователей интернет-сервисов, но и в объекты непрерывной генерации и обработки данных. В сочетании с описанными в статье рисками потенциального ущерба и возрастающим числом субъектов угроз это обусловливает острую потребность в формировании общегражданской культуры информационной безопасности.

СПИСОК ЛИТЕРАТУРЫ

1. Щербаков А.Ю. Четвертый факультет Высшей школы КГБ в зеркале истории великой страны // Вестник современных цифровых технологий. 2024. №20. С. 6-10. EDN: XCQVUO.

2. Дмитриева Е.В. Эволюция представления компьютерной терминологии // Тульский научный вестник. Серия История. Языкознание. 2021. №3(7). С.49-56. DOI: 10.22405/2712-8407-2021-3-49. EDN: AXZEHL.

3. Кузнецов А.В. Эволюция реагирования на инциденты информационной безопасности // Защита информации. Инсайд. 2024. №5(119). С.14-20. EDN: BHTBVA.

4. Исупов А.М., Мартышкин С.А., Прохоров Д. В. и др. Концептуальный анализ развития цифровизации в контексте новой цифровой экономики и по-литики // Вопросы национальных и федеративных отношений. 2022. Т.12. №7(88). С.2544-2555. DOI: 10.35775/PSI.2022.88.7.024. EDN: MNQSJZ.

5. Бандурин Г. И. 50 лет на передовой информационной отрасли // Автоматика, связь, информатика. 2020. №6. С.2-5. EDN: JYFPUU.

6. Никольская О. К. PROGRAMMA 101" фирмы "Оливетти": от социальной утопии к первому персональному компьютеру // Вестник современных цифровых технологий. 2020. №5. С.63-74. EDN: AELNTN.

7. Мещерякова А.Б. Эволюция цифрового доступа: переход от домашних компьютеров к мобильным технологиям в России // Вестник Академии знаний. 2024. №5(64). С.274-278. EDN: VHJSND.

8. Белозеров А.А., Вахлаков Д. В., Мельников С.Ю. и др. Технологические аспекты построения системы сбора и предобработки корпусов новостных текстов для создания моделей языка // Известия ЮФУ. Технические науки. 2016. №12(185). С.29-42. DOI: 10.18522/2311-3103-2016-12-2942 EDN: YFOLIH.

9. Кочкин А. В. Эволюция роли медиа в доктринах информационной безопасности США в XXI веке // Вестник Московского университета. Серия 10: Журналистика. 2024. Т.49. №5. С.100-114. DOI: 10.55959/msu.vestnik.journ.5.2024.100114. EDN: BTTCGU.

10. Ваничкина А.С. Трансформация ценностной картины мира эпохи Web 3.0 (на материале коммуникации участников рынка NFT-искусства) / Информационная безопасность и межкультурная коммуникация в контексте цифровой трансформации: Сборник научных трудов. М.: МГЛУ. 2022. С.267-276. EDN: HRGYEQ.

11. Альпидовская М.Л. Цифровой Левиафан // Вопросы политической экономии. 2021. №1. С.152-164. DOI: 10.5281/zenodo.4666342. EDN: ZSLKIT.

12. Розяева Т.Н., Шарашкина Т.П. Индустрия 4.0: эволюция взглядов и обеспечение информационной безопасности // Russian Economic Bulletin. 2021. Т.4. № 3. С.105-108. EDN: PDDDHP.

13. Гавдан Г. П., Горбатов В.С., Дураковский А.П., Наталичев Р.В. Эволюция и парадоксы нормативной базы обеспечения безопасности объектов критической информационной инфраструктуры // Безопасность информационных технологий. 2021. Т.28. №3. С.6-27. DOI: 10.26583/bit.2021.3.01. EDN: JIMDXU.

14. Абрамов В.И., Евдокимов Д.С. Автоматизированные системы управления экономикой СССР как прообразы современных ситуационных центров Российской Федерации // Региональные проблемы преобразования экономики. 2020. №7(117). С.13-24. DOI: 10.26726/1812-7096-2020-7-13-24. EDN: HEUVKA.

15. Былевский П.Г. Философско-культурологический анализ доверия к технологиям «искусственного интеллекта» // Вестник Кемеровского государственного университета культуры и искусств. 2024. №68. С.65-77. DOI: 10.31773/2078-1768-2024-68-65-77. EDN: EFACYO.

УДК: 004.056

Современные методы менеджмента информационной безопасности, основанные на риск-ориентированном подходе

I.A. Beloshitsky, M.Yu. Tolstykh

Modern Methods of Information Security Management Based on a Risk-Oriented Approach

Abstract. *The article is devoted to an overview of modern approaches to cyber risk management with an emphasis on the realities in the field of information security in the Russian Federation. Qualitative, quantitative and hybrid approaches to risk assessment, their methodological foundations, advantages and limitations are considered in detail. Particular attention is paid to modern cyber threats relevant to Russian organizations, such as social engineering and phishing, as well as new threats associated with the use of artificial intelligence technologies (fraud using deepfakes). In conclusion, practical recommendations are offered for increasing cyber resilience through information security management of organizations based on a risk-oriented approach.*

Keywords: *information security, information security management, risk management, threat, cyber risk.*

И.А. Белошицкий [1]
М.Ю. Толстых [2]
[1]*Магистрант,*
Московский государственный лингвистический университет.
E-mail: ibeloschitsky@gmail.com
[2]*Кандидат технических наук, доцент кафедры международной информационной безопасности,*
Московский государственный лингвистический университет.
E-mail: marina_lion@mail.ru

Аннотация. *Статья посвящена обзору современных подходов к управлению киберрисками с акцентом на реалии в сфере защиты информации Российской Федерации. Подробно рассмотрены качественный, количественный и гибридный подходы к оценке рисков, их методологические основы, преимущества и ограничения. Особое внимание уделено современным киберугрозам, актуальным для российских организаций, таким как социальная инженерия и фишинг, а также новым угрозам, связанным с применением технологий искусственного интеллекта (мошенничество с помощью дипфейков). В заключение предложены практические рекомендации для повышения киберустойчивости посредством менеджмента информационной безопасности организаций на базе риск-ориентированного подхода.*

Ключевые слова: *информационная безопасность, менеджмент информационной безопасности, управление рисками, угроза, киберриск.*

ВВЕДЕНИЕ

В условиях цифровизации и усложнения ИТ-инфраструктуры российские организации сталкиваются с постоянно растущим спектром киберугроз. Фишинговые атаки, программы-вымогатели (ransomware), утечки данных и другие инциденты информационной безопасности (ИБ) значительно участились за последние годы. Например, в 2023 году в доменной зоне .RU было заблокировано более чем в 5 раз больше фишинговых и скам-ссылок по сравнению с 2022 годом[1]. При этом, как свидетельствует аналитика, социальная инженерия остается одним из ключевых методов компрометации: почти в половине успешных кибератак на организации в 2024 году злоумышленники использовали методы обмана пользователей[2].

Одновременно возрастает сложность и масштабность атак: киберпреступники применяют инструменты искусственного интеллекта для автоматизации фишинга и создания достоверного поддельного контента, что порождает новые угрозы. Так, учащаются случаи использования deepfake-технологий в мошеннических схемах (генерация фальшивых голосов и видео) [1].

Атаки типа ransomware продолжают наносить серьезный ущерб бизнесу. Согласно данным Group-IB[3], в 2022-м году число атак вымогателей на рос-

[1] Лаборатория Касперского. Число заблокированных фишинговых ссылок в России выросло в 5 раз: пресс-релиз от 22 января 2024 г. – URL: https://www.kaspersky.ru/about/press-releases (Дата обращения: 20.06.2025).

[2] Positive Technologies. Актуальные киберугрозы для организаций: итоги 2024 года: аналитический отчет. – URL: https://ptsecurity.com/ru-ru/research/analytics (Дата обращения: 20.06.2025).

[3] Group-IB. Количество кибератак хакеров-вымогателей в России в 2022 году выросло в три раза // ТАСС, 17 марта 2023 года. – URL: https://tass.ru (Дата обращения: 24.06.2025).

сийские компании выросло почти втрое относительно 2021-го года, и на ransomware пришлось примерно 68% всех расследованных инцидентов. Отдельные группировки требовали рекордные суммы выкупа до 1 млрд руб., как в случае с группой OldGremlin. Средний простой в работе атакованной организации тогда достигал 14 дней.

В 2023 году тенденция сохранилась: доля шифровальщиков составила 57% от всего вредоносного ПО, использованного в успешных атаках, причем многие преступные группы практиковали двойное вымогательство (одновременное шифрование данных и утечка украденного). Ущерб от утечек конфиденциальных данных также растет, это проявляется в ужесточении штрафов и усилении репутационных потерь для российских компаний ввиду инцидентов, произошедших с персональными данными[4].

В таких условиях эффективное управление рисками ИБ становится ключевым элементом обеспечения устойчивости организации. Риск ИБ можно определить как вероятность реализации угрозы, умноженную на величину потенциального ущерба [2, 3]. В отечественной нормативной базе, например, в ГОСТ Р 51897-2021 (эквивалент ISO Guide 73), применяется подход, при котором риск определен как комбинация вероятности наступления события и последствий.

Управление ИБ-рисками представляет собой непрерывный процесс выявления, оценки и обработки рисков, позволяющий сфокусировать ресурсы защиты на наиболее критичных направлениях [4, 5]. Это особенно важно, поскольку ресурсы кибербезопасности всегда ограничены, а потенциальных угроз больше, чем возможностей их полностью нейтрализовать. Невозможно защитить абсолютно все активы от всех видов атак, поэтому требуется расстановка приоритетов, чему способствует систематическая оценка и менеджмент рисков.

ОСНОВНЫЕ ПОДХОДЫ К ОЦЕНКЕ И УПРАВЛЕНИЮ РИСКАМИ ИНФОРМАЦИОННОЙ БЕЗОПАСНОСТИ

Качественный подход: методологические основы, преимущества и ограничения

Качественный анализ рисков основывается на экспертизе и использовании описательных категорий для оценки вероятности и последствий потенциальных угроз. В рамках этого подхода риски ранжируются вербально, например, как «низкий», «средний» или «высокий», т.е. без прямого перевода оценок в денежные величины. Оценка проводится методом экспертного обсуждения, интервью и анкетирования с опорой на знания специалистов об активах, известных уязвимостях и возможных сценариях атак.

Достоинство подхода – относительная простота и быстрота внедрения: не требуются большие массивы статистических данных или сложные вычисления, достаточно привлечения компетентных экспертов. Качественный метод особенно полезен на начальных этапах построения системы менеджмента рисков или при ограниченных ресурсах организации и позволяет оперативно получить общее представление о профиле рисков, выявить области наибольшей уязвимости.

Существуют различные методики, реализующие качественный подход. Примеры:

- **OCTAVE** (от англ. «Operationally Critical Threat, Asset, and Vulnerability Evaluation» – оценка критически важных для эксплуатации угроз, активов и уязвимостей) ориентирована на анализ рисков силами самой организации;
- **STRIDE** (от англ. «spoofing, tampering, repudiation, information disclosure, denial of service, elevation of privilege» – подмена личности, модификация данных, отказ от ответственности, разглашение информации, отказ в обслуживании, повышение привилегий) – разработанная Microsoft методика, применяемая для анализа угроз в архитектуре программного обеспечения;
- **FRAP** (от англ. «Facilitated Risk Analysis Process» – упрощенный процесс анализа рисков) – фокусируется на коллективной быстрой оценке только самых критичных рисков без детального рассмотрения малозначимых угроз.

Общим для этих методик является опора на экспертное выявление и описание наиболее значимых для бизнеса сценариев киберрисков.

Преимущества качественного подхода заключаются в том, что он не требует сложных вычислений или специализированных инструментов: его относительно легко адаптировать под конкретную отрасль и организацию, а первые результаты можно получить в весьма короткие сроки.

Недостатки же связаны с субъективностью оценок и ограниченной доказательностью выводов в количественном выражении. Поскольку шкалы носят описательный характер, итоговые рейтинги риска зависят от мнения экспертов и могут различаться при смене команды оценщиков. Кроме того,

[4] Tadviser. Штрафы за утечку данных в России. – URL: https://www.tadviser.ru (Дата обращения: 20.06.2025).

качественные оценки трудно напрямую соотнести с финансовыми показателями: не имея точных цифр, сложно обосновать руководству размер бюджета, необходимого для снижения того или иного риска.

Качественный анализ не дает точной оценки потенциального ущерба и не позволяет количественно измерить эффективность мер защиты. Тем не менее, методы качественной оценки остаются важным инструментом управления ИБ-рисками, особенно когда нужно быстро охватить широкий круг рисков или отсутствуют данные для построения детальной количественной модели. В технологически зрелых организациях целесообразно дополнять качественную оценку более точными расчетами по ключевым рискам, когда это допустимо.

Количественный подход: методологические основы, преимущества и ограничения

Количественный анализ рисков предполагает использование числовых показателей, моделей и формул для оценки вероятности наступления негативных событий и ожидаемого ущерба от них. В этом подходе риск выражается в количественном виде как произведение вероятности инцидента на величину его последствий. Классическим примером служит метрика ALE (от англ. «Annual Loss Expectancy» – ожидаемый годовой убыток), вычисляемая как произведение единовременного ущерба от инцидента на ожидаемую частоту такого инцидента в год.

Количественный подход переводит абстрактные угрозы на понятный бизнесу язык денег. Результаты представляются в денежных единицах или других конкретных метриках, что позволяет напрямую оценивать экономическую значимость рисков и эффективность инвестиций в защитные меры.

Для реализации количественной оценки обычно требуется собрать обширные данные о прошлых инцидентах, статистику срабатывания угроз, вероятностные характеристики уязвимостей и т.д. На основе указанной информации строятся математические модели риска. Применяются, в частности, методы актуарных расчетов и статистического моделирования.

Например, **методология FAIR** (от англ. «Factor Analysis of Information Risk» – факторный анализ информационных рисков) задает формальную модель расчета вероятности и ущерба по каждому риску; **анализ ALE**, упомянутый выше, позволяет оценить средние годовые потери; **метод Монте-Карло** предполагает компьютерную имитацию множества сценариев для получения распределения возможных убытков. Такие методы дают более точную количественную картину совокупного киберриска, что особенно актуально для крупных организаций, где управление ИБ-рисками интегрировано в общий процесс корпоративного риск-менеджмента.

Преимущества количественного подхода состоят в его объективности и ориентации на бизнес-метрики. Решения в области безопасности можно обосновывать языком цифр: сравнивая рассчитанные величины рисков с затратами на контрмеры, руководство получает возможность оценить эффективность внедрения защитных мер.

Количественные оценки также позволяют ранжировать риски и системы по критичности в денежном выражении и соотнести приоритизацию мер защиты с общей финансово-экономической логикой компании.

Недостатки использования количественных методов связаны прежде всего с высокими требованиями к данным и ресурсам. Для корректной оценки необходимы достоверные статистические показатели: частоты инцидентов, распределения убытков, вероятности различных событий, сбор и актуализация такой информации трудоемки. Если эмпирических данных недостаточно, результаты моделей могут быть ненадежными.

Количественные методы сложны во внедрении: требуются квалифицированные аналитики, специальные программные инструменты, значительное время на моделирование и верификацию. Дополнительное ограничение – трудность оценки новых и редко встречающихся угроз. А в ситуации, когда организация сталкивается с беспрецедентными рисками (например, атаками типа APT (от англ. Advanced Persistent Threat – целевая, долгосрочная и скрытая кибератака) или уязвимостями нулевого дня), может не оказаться данных для расчетов, количественная оценка в таких случаях либо невозможна, либо имеет большой разброс неопределенности.

Несмотря на сложности в реализации, количественный анализ предоставляет ценные возможности для продвинутого управления киберрисками. Он незаменим, когда требуется детальное экономическое обоснование мер безопасности или построение моделей ущерба для различных сценариев. На практике переход к количественным методам часто связан с ростом зрелости процессов: как правило, организация начинает с качественной приоритизации рисков, а по мере накопления данных и повышения требований бизнеса постепенно внедряет более точные количественные модели.

Гибридный подход. Методологические основы, преимущества и ограничения

Гибридный (смешанный) подход сочетает элементы качественного и количественного анализа рисков. Он выступает своеобразным компромиссом, позволяющим воспользоваться сильными сторонами обоих методов и смягчить их недостатки. На практике это может реализовываться либо параллельным использованием разных шкал оценивания, либо последовательным применением методов на разных этапах.

Например, изначально каждому риску могут присваиваться баллы (допустим, по шкале 1–5) с последующим переводом этих баллов в категориальные уровни риска («низкий», «средний», «высокий»). Обратный вариант: сначала проводится вербальная оценка рисков, которая позже калибруется количественно.

Преимущества гибридного подхода – его гибкость и эффективность. Качественная составляющая позволяет быстро и экономично выявить проблемные области и отсечь второстепенные риски, не распыляя ресурсы. Количественная составляющая дает точные расчеты там, где это действительно необходимо для принятия решений (например, по рискам с потенциально катастрофическими последствиями). Комбинированный подход также может ускорять процесс оценивания: по некоторым данным, полный цикл оценки киберрисков при комбинированном методе проходит в разы быстрее, чем при использовании только количественного анализа.

Недостатки гибридного подхода во многом аналогичны недостаткам его компонентов: сохраняется элемент субъективности на качественных этапах и высокая трудоемкость на количественных. Важно правильно выстроить методику, чтобы переход от качественной оценки к количественной был обоснованным и опирался на результаты первого этапа; иначе две разные оценки могут дать разрозненные выводы.

Краткое сравнение подходов к оценке рисков

Многие эксперты отмечают [6], что на сегодняшний день наиболее распространены именно качественные и смешанные подходы к оценке киберрисков, тогда как полностью количественный подход применяется сравнительно реже. В таблице 1 приведены основные характеристики анализируемых методов оценки ИБ-рисков.

Таблица 1

Сравнительный анализ подходов к оценке рисков ИБ

Характеристика	Качественный подход	Количественный подход	Гибридный подход
Простота внедрения	Высокая	Низкая	Средняя
Необходимость сбора данных	Низкая	Высокая	Средняя
Объективность оценки	Низкая	Высокая	Средняя
Точность прогнозирования ущерба	Низкая	Высокая	Средняя
Использование экспертных оценок	Да	Частично	Да
Сложность расчетов	Низкая	Высокая	Средняя

В целом постановка вопроса о том, какой подход лучше – качественный или количественный, не имеет смысла, поскольку в современном понимании необходимо комбинировать оба подхода в зависимости от контекста и доступных данных. В арсенале менеджера по ИБ (офицера безопасности) должны присутствовать разные инструменты оценки. Гибридный подход предоставляет такую возможность: использовать быстрые экспертные оценки для широкого охвата множества рисков и точные количественные расчеты – для узкого круга наиболее критичных из них. Практика показывает, что именно разумное сочетание методов позволяет достичь оптимального баланса между оперативностью оценки и точностью результатов.

СОВРЕМЕННЫЕ КИБЕРУГРОЗЫ И ВЫЗОВЫ

Эффективное управление ИБ-рисками невозможно без учета актуального ландшафта угроз [7]. Ниже рассмотрены три группы кибервызовов, на наш взгляд наиболее актуальные для российских организаций, а также специфические аспекты работы с соответствующими рисками.

Фишинг по-прежнему является самым массовым видом кибератак, нацеленным на человеческий фак-

тор. Посредством электронных писем, сообщений в мессенджерах или социальных сетях, телефонных звонков злоумышленники вынуждают жертв добровольно раскрыть конфиденциальные данные или совершить опасные действия. Цели «фишеров» становятся разнообразнее: теперь это не только учетные данные от корпоративных систем и платежных сервисов, но и аккаунты в мессенджерах.

Компрометация аккаунта в WhatsApp или Telegram часто используется как первый шаг многоступенчатого мошенничества: захватив аккаунт, злоумышленники пытаются атаковать доверенных контактных лиц жертвы [8]. Для рассылки фишинговых сообщений по-прежнему доминирует e-mail (оценочно до 90% случаев), однако заметно растет доля схем с использованием SMS, социальных сетей и мессенджеров. Технические приемы также эволюционируют: помимо традиционных вложений-троянов (офисные документы, архивы), злоумышленники активно используют ссылки на файлообменники и облачные диски для загрузки вредоносного ПО, а также легитимные онлайн-сервисы (например, Google Docs) для размещения фишинговых страниц.

Особую опасность представляют таргетированные фишинговые атаки и компрометация бизнес-переписки. В качестве примера можно упомянуть кампанию 2023-го года группировки Red Wolf, нацеленную на российские промышленные предприятия [9]. Атакующие тщательно изучали внутренних адресатов, писали письма с учетом отраслевой специфики, имитировали переписку от имени контрагентов или госорганов. В ряде случаев такой целевой фишинг был лишь этапом сложной комбинированной атаки: получив начальный доступ через письма, злоумышленники разворачивали в сети жертвы шпионские программы, похищали данные, а затем использовали их для вымогательства или последующих разрушительных действий.

По совокупным отчетам [10], около 45% всех успешных кибератак на организации в мире и в РФ в 2023-м году реализованы с применением ***методов социальной инженерии***. Это подчеркивает, что «человеческий фактор» остается критической зоной риска, требующей особого внимания.

Для снижения риска подобных кибератак российские компании внедряют программы обучения сотрудников основам кибергигиены, проводят имитационные фишинг-рассылки (тренинги), используют технологии фильтрации фишинговых писем и сообщений, внедряют системы обнаружения компрометации корпоративных доменов (такие как DMARC, от англ. «Domain-based Message Authentication, Reporting and Conformance» – техническая спецификация для снижения количества спамовых и фишинговых электронных писем). Тем не менее, инциденты показывают, что просветительские меры эффективны лишь отчасти: злоумышленники постоянно совершенствуют психологические уловки, эксплуатируя доверие к известным брендам, чувство страха или срочности, изобилие правдоподобных деталей.

Например, в 2025 году в России была выявлена новая схема[5]: мошенники звонят под видом сотрудников социологического опроса, чтобы незаметно записать голос жертвы, а затем использовать эту запись для реалистичных голосовых фишинг-звонков с требованием перевести деньги. Подобные тенденции требуют от организаций непрерывного мониторинга актуальных сценариев атак и регулярного обновления как технических средств защиты (почтовых фильтров, антиспам/антифишинг систем), так и программ повышения осведомленности персонала.

Вымогательские атаки с шифрованием данных (ransomware) остаются одним из самых критичных киберрисков для бизнеса. В последние годы изменился и масштаб проблемы, и подходы самих злоумышленников. Если ранее группы ransomware в основном нацеливались на зарубежные компании, то с 2020-2022 годов фиксируется резкий рост атак на российский бизнес. Эксперты связывают это с несколькими факторами: утечкой в открытый доступ исходных кодов известных вымогательских программ (Conti, LockBit), появлением на теневом рынке сервисов Ransomware-as-a-Service (RaaS), а также общей «коммерциализацией» киберпреступности [11].

По данным лаборатории цифровой криминалистики Group-IB[6], в 2022 году в России число атак с применением шифровальщиков увеличилось на в 3 раза по сравнению с предыдущим годом. На программы-вымогатели пришлось до 68% всех серьезных инцидентов, расследованных специалистами. Чаще всего жертвами становились компании розничной торговли, промышленности и страхового сектора, вероятно, из-за относительно более слабой защищенности этих отраслей и их готовности платить за восстановление работы. К 2023 году атаки вымогателей охватили практически все секторы экономики: по данным Positive Technologies[7], рост числа таких атак отмечен во всех отраслях без исключения.

[5] РИА Новости: Эксперт объявил, зачем аферисты звонят россиянам под видом соцопросов. – URL: https://ria.ru/20250611/moshenniki-2022156038.html (Дата обращения: 24.06.2025).

[6] Group-IB. Количество кибератак хакеров-вымогателей в России в 2022 году выросло в три раза. – URL: https://tass.ru (Дата обращения: 24.06.2025).

[7] Positive Technologies. Актуальные киберугрозы для организаций: итоги 2023 года: аналитический отчет. – М.: Positive Technologies, 2024. – 48 с.

Злоумышленники усовершенствовали тактику от простого шифрования к двойному вымогательству: перед тем как зашифровать базы данных, они похищают значительный объем конфиденциальной информации. Затем жертве предъявляется двойное требование выкупа: не только за ключ дешифрования, но и за обещание не публиковать украденные данные. Официально выплачивать выкуп не рекомендуется (а в некоторых странах прямо запрещено законом), однако многие организации – особенно из критической инфраструктуры – оказываются перед тяжелой дилеммой, когда остановка бизнес-процессов грозит несоизмеримо большими потерями.

Управление риском в отношении атаки ransomware требует комплексных мер. Необходимо уделять внимание проактивной защите: регулярно резервировать и изолированно хранить критичные данные, сегментировать сеть, своевременно обновлять ПО для устранения уязвимостей, что снизит вероятность успешного проникновения и масштаб ущерба. Также необходима разработка детальных планов реагирования на инциденты вымогательства и планов обеспечения непрерывности бизнеса, чтобы в случае атаки быстро среагировать и минимизировать простои.

Дополнительный вызов – появление так называемых ***wiper-атак и псевдо-ransomware*** со стороны так называемых хактивистов (мотивация которых не финансовая, а идеологическая или деструктивная), особенно на фоне текущей геополитической обстановки. В таких случаях цель злоумышленников заключается не в получении выкупа, а в том, чтобы максимально уничтожить или вывести из строя ИТ-инфраструктуру жертвы, создать резонанс.

Это усложняет модель риска: даже системы, не содержащие критичных данных, могут быть атакованы из вредительских или саботажных побуждений, просто чтобы прервать деятельность организации. В ответ на подобные сценарии организации, отнесенные к объектам КИИ, по требованиям 187-ФЗ обязаны внедрять системы обнаружения и предотвращения атак (государственная система обнаружения, предупреждения и ликвидации последствий компьютерных атак на информационные ресурсы РФ, ГосСОПКА) и уведомлять регуляторов о каждом инциденте. Данные меры призваны повысить коллективную готовность противостоять деструктивным атакам.

Стандарты также учитывают угрозу ransomware. Так, ГОСТ Р 57580.3-2022 для финансовых организаций прямо требует учитывать риск реализации вымогательских киберугроз и включать соответствующие индикаторы (например, долю зашифрованных критичных данных при атаке, время простоя, эффективность восстановления из резервных копий) в систему КПУР. В целом, на фоне сохраняющейся высокой активности вымогателей, менеджерам по рискам следует планировать оперативные действия по минимизации ущерба, исходя из предположения, что атака в любом случае когда-то состоится.

Появление в последние годы доступных ***технологий искусственного интеллекта*** породило новую категорию киберрисков. С одной стороны, искусственный интеллект дает специалистам по безопасности мощные средства для автоматизации мониторинга и анализа инцидентов. С другой – эти же технологии вошли в арсенал злоумышленников, позволив существенно повысить масштабируемость и правдоподобие атак социальной инженерии. Один из наиболее обсуждаемых рисков – мошенническое ***использование deepfake*** (генерируемых нейросетью поддельных изображений, аудио или видео).

В 2023 году во всем мире зафиксирован всплеск инцидентов с применением дипфейков, Россия не стала исключением: летом 2023 года неизвестные хакеры взломали трансляции ряда региональных телеканалов и радиостанций, показав в эфире дипфейковое видео с экстренным обращением, имитирующим обращение Президента РФ[8]. В этом фальшивом сообщении объявлялось военное положение и звучали иные провокационные заявления; позднее власти официально опровергли эту информацию, признав факт взлома. Данный инцидент продемонстрировал потенциальную опасность deepfake-технологий как инструмента информационных атак и дезинформации.

В сфере финансового мошенничества возможности дипфейков также стремительно растут. Уже существуют сервисы, позволяющие сгенерировать достаточно правдоподобный голосовой дипфейк человека, имея лишь несколько десятков секунд записи его речи. Этим начали пользоваться телефонные мошенники. Описанная выше схема с «соцопросом» – лишь один пример: цель атакующих – записать голос жертвы, чтобы затем, изменив интонации и тембр с помощью нейросети, позвонить от имени ее знакомого или руководителя и вытянуть деньги.

Помимо дипфейков, злоумышленники применяют технологии искусственного интеллекта и для

[8] ТАСС: Эксперт: дипфейк с «обращением» Путина сделали при помощи технологии замены лица. – URL: https://tass.ru (Дата обращения: 24.06.2025).

других целей. ***Генерация реалистичных фишинговых текстов*** на русском языке без типичных ошибок и шаблонов, распознаваемых фильтрами, стала тривиальной задачей с появлением больших языковых моделей (пример, чат-боты наподобие ChatGPT). Автоматизация разведки по открытым источникам и подбор персонализированных уловок также упростились: достаточно написать скрипт с AI-модулем, чтобы проанализировать сотни профилей в соцсетях и составить убедительное мошенническое сообщение для каждой жертвы.

Дополнительная угроза – ***маскировка вредоносного ПО*** под популярные приложения на базе искусственного интеллекта. Пользователи охотно скачивают новые «чудо-нейросети», чем пользуются создатели вредоносного ПО. Например, в 2023 году обнаружена схема, когда на фишинговом сайте предлагалась загрузка некой программы «DeepSeek-R1», позиционируемой как новая популярная нейросеть для ПК. Под ее видом распространялся ранее неизвестный троян для кражи данных [12].

В другом случае злоумышленники создали поддельное мобильное приложение, якобы предоставляющее доступ к ChatGPT, и распространяли его через фишинговую рассылку, в результате сотни Android-устройств оказались заражены шпионским ПО. Такие инциденты показывают, что ажиотаж вокруг технологий сам по себе стал приманкой для атак. Это нужно учитывать при оценке рисков: всплеск интереса аудитории к новой AI-платформе почти гарантированно сопровождается появлением мошенников, пытающихся использовать его для собственной выгоды.

В целях противодействия AI-угрозам отечественные регуляторы и эксперты также предпринимают шаги. В 2023 году Роскомнадзор официально предупредил граждан об опасности телефонных звонков с применением генерации голоса, рекомендовав не доверять таким вызовам и использовать кодовые слова при финансовых операциях по телефону. В Государственной Думе РФ обсуждаются инициативы о введении ответственности за создание и распространение вредоносных дипфейков, особенно если они направлены против государственных институтов [13].

На уровне организаций методы управления рисками уже начинают включать сценарии атак с применением искусственного интеллекта: проводятся учебные бизнес-игры (table-top учения) по моделированию таких инцидентов, оценивается потенциальный ущерб от них, внедряются дополнительные контроли – например, технические средства обнаружения дипфейков по характерным артефактам генерации.

Безусловно, данная сфера считается относительно новой, и эффективных инструментов борьбы с угрозами подобного вида в настоящее время насчитывается не так много, но понимание значения и масштаба угроз и рисков стремительно развивается. Комбинация таких мер, как технологические фильтры (например, для распознавания синтетического голоса), обновленные политики безопасности (строгая верификация важных запросов по нескольким каналам) и повышенная осведомленность сотрудников, поможет снизить этот новый класс рисков до приемлемого уровня.

ЗАКЛЮЧЕНИЕ

Управление рисками ИБ в современных условиях является неотъемлемым элементом обеспечения киберустойчивости организаций. Для российского бизнеса и госструктур это особенно актуально: с одной стороны, регуляторы требуют внедрения риск-ориентированных процессов (исполнение требований законодательства о безопасности критической информационной инфраструктуры, стандартов Банка России, регламентов по защите данных), с другой – ландшафт угроз усложняется, появляются новые векторы атак.

Рассмотренные в статье подходы (качественный, количественный и гибридный) не противопоставляются друг другу, а напротив, взаимно дополняют. Качественные методы управление рисками ИБ позволяют быстро охватить широкий спектр рисков и выдвинуть первоочередные гипотезы о наиболее опасных сценариях. Количественные методы предоставляют инструментарий для глубокой аналитики и экономического обоснования инвестиций в безопасность.

Их комбинация на практике дает оптимальный баланс: оперативность и экономичность оценки при достаточной точности в критически важных вопросах.

На основании результатов исследования для повышения эффективности управления киберрисками в организации нами выработаны следующие рекомендации:

1. Интеграция в корпоративный менеджмент рисков. Процессы управления ИБ-рисками следует встроить в общую систему корпоративного риск-менеджмента, обеспечив внимание к ним на уровне высшего руководства. Киберриски должны рассматриваться как часть бизнес-рисков, влияющих на стратегические цели.

2. Следование праву, стандартам и развитие собственных метрик. Необходимо соблюдать требования и руководства национальных стандартов (например, ГОСТ Р 57580.3-2022) и регуляторов, но не ограничиваться формальной их имплементацией. Полезно развивать внутренние метрики, базы данных инцидентов и модели рисков, релевантные специфике конкретного бизнеса, поскольку это даст более точное понимание угроз и эффективности мер.

3. Регулярная актуализация модели угроз. Угрозы ИБ стремительно меняются (появляются новые уязвимости, методы атак с применением технологий искусственного интеллекта, учитываются геополитические факторы). Следует систематически пересматривать модель угроз и оценки рисков, используя актуальные данные об инцидентах и тенденциях. Такая динамическая оценка позволит не пропустить новые значимые риски.

4. Комбинирование методов оценки. Практично применять качественные оценки для скрининга и первичной приоритизации большого списка рисков, после чего наиболее существенные из них подвергать углубленному количественному анализу. Например, сначала экспертно выделить 10 самых актуальных рисков, а затем для каждого из них оценить потенциальный ущерб и вероятность на основе статистики (рассчитать ALE или провести моделирование), что обеспечит рациональное распределение ресурсов: детальный анализ только там, где это действительно важно.

5. Обучение персонала и культивирование киберкультуры. Человеческий фактор остается ключевым, поскольку многие инциденты происходят из-за ошибок или неосведомленности сотрудников. Необходимо регулярно обучать сотрудников всех уровней принципам кибербезопасности и информировать их о новых угрозах (фишинг-схемах, социальной инженерии и т.п.). Формирование устойчивой киберкультуры в коллективе заметно снизит риск успешных атак.

6. Планирование реагирования и устойчивости. Помимо предотвращения инцидентов, организация должна быть готова к тому, что серьезный инцидент все же произойдет. Разработанные и протестированные планы реагирования на киберинциденты и планы обеспечения непрерывности помогут минимизировать ущерб и время простоя. В контексте рисков типа ransomware или крупной утечки данных следует заранее продумать шаги: кого уведомлять, как изолировать системы, как восстановиться из резервных копий, нужны ли шаблоны публичных заявлений и т.д.

Реализация перечисленных мер, основанная на сочетании различных подходов к оценке рисков, позволит выстроить проактивную и адаптивную систему менеджмента ИБ-рисков. Система не только обеспечит более рациональное распределение ресурсов защиты, но и повысит общую готовность противостоять современным киберугрозам. В конечном итоге это укрепит устойчивость бизнеса и национальной цифровой инфраструктуры, снизив потенциальный ущерб от инцидентов и увеличив возможность быстрее восстанавливаться после них, накапливать полезный опыт и формировать лучшие практики управления ИБ.

СПИСОК ЛИТЕРАТУРЫ

1. Старостенко Н. И. Криминалистическое прогнозирование хищений, совершаемых с использованием «deepfake»-технологий // Вестник Сибирского юридического института МВД России. – 2023. – № 2(51). – С. 187-192.
2. Овчинникова Е. А. Основы управления рисками информационной безопасности. Планирование управления рисками: учебное пособие / Е. А. Овчинникова, В. А. Герасимов, А. Д. Воропаев. – Новосибирск: Сибирский государственный университет телекоммуникаций и информатики. 2024. – 81 с.
3. Веселов Г. Е. Менеджмент риска информационной безопасности: учебное пособие / Г. Е. Веселов, Е. С. Абрамов, А. К. Шилов. – Ростов-на-Дону: ЮФУ, 2016. – 107 с.
4. Милославская Н. Г. Управление рисками информационной безопасности. Учебное пособие для вузов / Н.Г. Милославская, М.Ю. Сенаторов, А.И. Толстой. – Москва: Горячая Линия–Телеком, 2024. – 130 с.
5. Астахов А. М. Искусство управления информационными рисками / А. М. Астахов. – 2-е изд. – Москва: ДМК Пресс, 2018. – 312 с.
6. Ларина О. И. К вопросу о развитии методологии идентификации киберриска / О. И. Ларина, Н. В. Морыженкова // Банковское дело. – 2023. – № 1. – С. 66-71.
7. Минаков А. В. Оценка модели рисков информационной безопасности: характеристика, проблемы и перспективы / А. В. Минаков // Экономика и бизнес: теория и практика. – 2023. – № 10-2(104). – С. 63-69.
8. Ренессанс вымогателей и мошенников // Системный администратор. – 2025. – № 1-2(266-267). – С. 24-38.

9. Токарев Е. В. Угрозы кибербезопасности в 2022 и 2023 году: различия, сходства и прогнозы на 2024 год / Е. В. Токарев // Научный аспект. – 2024. – Т. 10, № 5. – С. 1239-1247.
10. Токарев Е. В. Угрозы кибербезопасности в 2023 году: анализ проблем и возможностей / Е. В. Токарев // Научный аспект. – 2023. – Т. 28, № 11. – С. 3431-3439.
11. Железко С. В. Используемая злоумышленниками бизнес-модель для проведения кибератак / С. В. Железко // Управление информационными ресурсами: Материалы XX Международной научно-практической конференции, Минск, 29 марта 2024 года. – Минск: Академия управления при Президенте Республики Беларусь, 2024. – С. 429-431.
12. Маршалл М. DeepSeek R1 и OpenAI Deep Research дали новое направление развития ИИ / М. Маршалл // БИТ. Бизнес & Информационные технологии. – 2025. – № 1(144). – С. 16-19.
13. Перина А. С. Квалификация цифровых преступлений против личности: проблемные вопросы / А. С. Перина // Вестник Югорского государственного университета. – 2023. – № 2(69). – С. 89-104.

УДК: 004.056

Статистические методы оценки криптостойкости генераторов псевдослучайных чисел

S. A. Mirzoyan

Statistical Methods for Assessing the Cryptographic Strength of Pseudorandom Number Generators

Abstract. *This paper explores statistical methods for assessing the cryptographic strength of pseudorandom number generators (PRNGs), with a focus on the application of the chi-square test to analyze the randomness of output sequences. The study aims to formalize the concept of cryptographic strength as a verifiable statistical hypothesis. Experimental results are presented for various types of PRNGs using the proposed approach. The effectiveness of the chi-square test in detecting hidden dependencies between sequence elements is demonstrated, supporting its use as a tool for evaluating cryptographic security.*

Keywords: *cryptographic strength, pseudorandom number generator, randomness testing, statistical analysis, chi-square test, bitstream.*

С. А. Мирзоян

Аспирант, преподаватель кафедры международной информационной безопасности, Московский государственный лингвистический университет.
E-mail: sergey.mirzoyan@bk.ru

Аннотация. *В статье рассматриваются статистические методы оценки криптостойкости генераторов псевдослучайных чисел (ГПСЧ). Особое внимание уделяется применению теста хи-квадрат для анализа случайности выходных последовательностей. Исследование направлено на формализацию понятия криптостойкости как проверяемой статистической гипотезы. Представлены результаты экспериментального анализа различных типов ГПСЧ с использованием предложенного подхода. Показана эффективность теста хи-квадрат в выявлении скрытых зависимостей между элементами последовательности, что позволяет использовать его в качестве инструмента оценки криптографической безопасности.*

Ключевые слова: *криптостойкость, генератор псевдослучайных чисел, тестирование случайности, статистический анализ, тест хи-квадрат, битовая последовательность.*

ВВЕДЕНИЕ

В современной криптографии безопасность информации напрямую зависит от качества используемых генераторов псевдослучайных чисел (ГПСЧ) [1]. Их применение распространяется на широкий круг задач: от формирования криптографических ключей до обеспечения целостности протоколов аутентификации [2]. От надёжности генератора зависит уровень защиты всей системы от возможных атак, включая восстановление ключевой информации и предсказание случайных параметров [3].

Криптостойкость ГПСЧ определяется его способностью вырабатывать последовательности, которые невозможно предсказать или восстановить без знания внутреннего состояния генератора. При этом недостаточно просто соответствовать требованиям равномерного распределения или отсутствия явных корреляций — необходимо также учитывать возможность формального доказательства случайности, основанного на статистических гипотезах [4].

Одним из наиболее популярных подходов к оценке случайности является использование статистических тестов, таких как тест хи-квадрат, тесты NIST SP 800-22 и набор DIEHARD. Эти методы позволяют формализовать процесс анализа и дать количественную оценку вероятности случайности последовательности. Особенно важным является использование теста хи-квадрат, который позволяет проверить согласованность эмпирического распределения частот с теоретическим законом равномерного распределения [5].

В данной работе делается акцент на том, что криптостойкость можно формально интерпретировать как статистическую гипотезу о случайности последовательности, и её можно оценить с помощью строгих математических методов. Это позволяет отказаться от субъективных экспертных оценок и перейти к объективному, воспроизводимому анализу.

Целью представленного исследования является разработка и обоснование методики оценки криптостойкости генераторов псевдослучайных чисел на основе статистического анализа битовых последовательностей. К задачам относятся:

- анализ существующих статистических тестов, применимых к оценке случайности;
- разработка алгоритма формальной проверки гипотезы о случайности на основе теста хи-квадрат;
- экспериментальная проверка эффективности предложенного подхода;

• сравнение результатов с другими стандартными методами.

ТЕОРЕТИЧЕСКИЕ ОСНОВЫ ГЕНЕРАТОРОВ ПСЕВДОСЛУЧАЙНЫХ ЧИСЕЛ

Генераторы псевдослучайных чисел (ГПСЧ) представляют собой алгоритмы, предназначенные для формирования последовательностей чисел, которые имитируют случайное поведение. Они находят широкое применение в различных областях информационных технологий, особенно в криптографии, где надёжность системы во многом зависит от качества используемого ГПСЧ [6].

В данном разделе рассматриваются основные типы ГПСЧ, их свойства, а также требования, которым они должны удовлетворять для обеспечения криптографической безопасности.

Классификация ГПСЧ

В зависимости от принципа работы и источника случайности, все генераторы можно разделить на два больших класса: детерминированные и недетерминированные.

Детерминированные ГПСЧ

Эти генераторы вырабатывают псевдослучайные последовательности на основе заданного начального значения — зерна (seed). Поскольку алгоритм работы полностью детерминирован, повторное использование одного и того же зерна приводит к воспроизводству точно такой же последовательности. Примеры таких генераторов:

- Линейный конгруэнтный генератор (LCG)
- Mersenne Twister [7]
- HMAC_DRBG

Детерминированные ГПСЧ широко используются в программном обеспечении благодаря высокой скорости работы и возможности точного воспроизведения результатов. Однако их предсказуемость делает их уязвимыми при использовании в криптографических целях, если значение зерна становится известным злоумышленнику [8].

Недетерминированные ГПСЧ

Такие генераторы строятся на основе физических процессов, например, теплового шума, радиоактивного распада или пользовательского ввода. Эти процессы по своей природе случайны и недетерминированы, что позволяет получать истинно случайные последовательности. Примеры:

- Аппаратные генераторы, использующие термальный шум
- Квантовые генераторы случайных чисел
- Генераторы, использующие флуктуации времени между событиями ввода-вывода

Недетерминированные ГПСЧ обеспечивают более высокую степень безопасности, но имеют ограничения по скорости и сложности реализации. Поэтому они чаще всего применяются в системах, где требуется максимальная криптостойкость [9].

Криптографически стойкие ГПСЧ (cryptographically secure pseudorandom number generator, CSPRNG)

Особый класс генераторов, сочетающий преимущества детерминированных и недетерминированных моделей. CSPRNG используют внешние источники энтропии для инициализации внутреннего состояния, после чего продолжают работу по детерминированному алгоритму [10]. При этом любое изменение состояния должно быть скрытым от наблюдателя, а выходная последовательность — непредсказуемой.

Примеры криптографически стойких генераторов:

- ChaCha20 RNG
- AES-CTR_DRBG
- Fortuna (разработан Брюсом Шнайером) [11]

Криптографически стойкие ГПСЧ играют ключевую роль в защите информации, поскольку обеспечивают баланс между производительностью и уровнем защиты [12].

Основные требования к криптостойкому ГПСЧ

Для того чтобы ГПСЧ мог быть использован в криптографических целях, он должен удовлетворять следующим основным требованиям:

- **Непредсказуемость.** Выходная последовательность не должна допускать предсказания следующего бита даже при наличии всей предыдущей истории. Это является ключевым свойством криптографической безопасности.
- **Необратимость.** Даже если часть последовательности стала известна, невозможно восстановить её предыдущие значения или текущее состояние генератора.
- **Равномерное распределение.** Биты выходной последовательности должны иметь равномерное распределение вероятностей: частота появления нулей и единиц должна стремиться к 50%.
- **Долгий период.** Период генератора — это длина последовательности, после которой она начинает повторяться. Для криптографического применения период должен быть достаточно большим, чтобы исключить возможность его перебора.
- **Устойчивость к атакам.** ГПСЧ должен быть спроектирован таким образом, чтобы противостоять различным типам атак, включая корреляционные, дифференциальные и временные атаки.

Угрозы и атаки на ГПСЧ

Несмотря на кажущуюся простоту использования, ГПСЧ подвержены ряду атак, которые могут компрометировать всю систему безопасности.

Предсказание следующего бита

Если злоумышленник может предсказать следующий бит последовательности, это даёт ему возможность восстановить ключи, пароли или другие секретные данные.

Восстановление внутреннего состояния

Многие ГПСЧ работают на основе внутреннего состояния, которое меняется со временем. Если злоумышленнику удастся восстановить это состояние, он сможет предсказывать все будущие значения.

Корреляционные атаки

Эти атаки основаны на анализе корреляций между различными частями последовательности. Особенно эффективны против линейных и слабо нелинейных генераторов.

Атаки на основе малой энтропии

Если источник энтропии (например, зерно) имеет низкое качество, злоумышленник может провести полный перебор возможных начальных состояний.

Атаки на реализацию

Иногда уязвимости возникают не в самом алгоритме ГПСЧ, а в его программной или аппаратной реализации. Например, ошибки в генерации зерна могут привести к предсказуемости.

Таким образом, выбор подходящего ГПСЧ зависит от задачи, в которой он будет применяться. Для криптографических целей необходимо использовать только криптографически безопасные генераторы, обладающие непредсказуемостью, необратимостью и устойчивостью к атакам.

В следующем разделе будут рассмотрены методы статистического тестирования, позволяющие формально проверить выполнение этих требований.

СТАТИСТИЧЕСКИЕ ТЕСТЫ ДЛЯ АНАЛИЗА СЛУЧАЙНОСТИ

Для оценки качества генераторов псевдослучайных чисел широко применяются методы математической статистики, позволяющие формально проверить гипотезу о случайности последовательности [13]. Эти методы используются как для первичной оценки свойств ГПСЧ, так и для проверки их соответствия требованиям криптографической безопасности.

Одним из наиболее популярных подходов является использование теста хи-квадрат (chi-square test), который позволяет оценить согласованность эмпирического распределения частот с теоретическим законом равномерного распределения. В данном разделе рассматриваются основные принципы статистического тестирования, описание теста хи-квадрат, его применение к битовым последовательностям и интерпретация результатов как меры случайности.

Общие принципы статистического тестирования

Статистическое тестирование ГПСЧ заключается в проверке гипотезы о том, что выходная последовательность соответствует случайному распределению. Для этого используется ряд математических процедур, которые выявляют отклонения от идеального поведения.

Основные этапы статистического тестирования:

• Формулировка нулевой гипотезы: предполагается, что последовательность случайна.

• Выбор статистики теста: определяется метрика, по которой будет оцениваться случайность.

• Расчёт p-значения: вероятность того, что наблюдаемое отклонение могло возникнуть случайно при условии справедливости нулевой гипотезы.

• Принятие или отвержение гипотезы на основе заданного уровня значимости (обычно 0.05).

Если p-значение меньше порогового уровня, то гипотеза о случайности отвергается, и последовательность считается неслучайной.

Тест хи-квадрат

Тест хи-квадрат — это классический метод математической статистики, используемый для сравнения наблюдаемых и ожидаемых частот в дискретных распределениях.

Формула теста:

$$\chi^2 = \sum_{i=1}^{k} \frac{(O^i - E^i)^2}{E^i}$$

где:

• O^i — наблюдаемая частота i-го значения,

• E^i — ожидаемая частота i-го значения,

• k — количество возможных исходов.

Интерпретация:

• Чем ближе значение χ^2 к нулю, тем выше согласованность между наблюдаемыми и ожидаемыми данными.

• По значению χ^2 вычисляется p-значение, которое сравнивается с уровнем значимости.

В случае бинарной последовательности, состоящей из нулей и единиц, можно проверить, насколько равномерно они распределены.

Пример

Рассмотрим последовательность длиной 1000 бит, где встречается 510 нулей и 490 единиц.

$$\chi^2 = \frac{(510-500)^2}{500} + \frac{(490-500)^2}{500} = \\ = \frac{100}{500} + \frac{100}{500} = 0.4$$

По таблицам распределения хи-квадрат с одной степенью свободы, это даёт p-значение ≈ 0.527, что больше 0.05. Следовательно, гипотеза о случайности принимается.

Применение теста хи-квадрат для оценки криптостойкости

Так как криптостойкий ГПСЧ должен обеспечивать равномерное распределение битов и отсутствие явных зависимостей между элементами, тест хи-квадрат может служить важным инструментом для формальной проверки этих свойств.

Основные направления применения:

1. Анализ частот появления битов.

Проверяется, насколько часто встречаются нули и единицы. Отклонение от 50% указывает на предсказуемость.

2. Выявление зависимостей между соседними элементами

Можно строить двумерную таблицу сопряжённости для пар битов (00, 01, 10, 11) и проверять, насколько равномерно они распределены.

3. Анализ окон фиксированного размера

Последовательность разбивается на блоки (например, по 8 бит), и для каждого блока подсчитывается частота его появления. Это позволяет обнаруживать повторяющиеся паттерны.

Преимущества использования теста хи-квадрат:

- простота реализации;
- возможность автоматизации;
- объективная количественная оценка случайности;
- широкая применимость к различным типам данных.

Ограничения:

- не учитывает сложные временные зависимости;
- может не обнаружить тонкие корреляции между элементами;
- требует достаточной длины последовательности для достоверности результатов.

Другие статистические тесты

Хотя тест хи-квадрат является мощным инструментом, он не является единственным методом статистического анализа ГПСЧ.

NIST SP 800-22

Это официальный набор тестов, рекомендованный Национальным институтом стандартов и технологий США для оценки случайности. Он включает такие тесты, как:

- частотный тест,
- тест последовательностей,
- тест длинных серий,
- тест блоков,
- тест ранга матрицы и др.

DIEHARD

Этот набор тестов был разработан Джорджем Марсальей и включает более 15 различных методов проверки случайности, включая анализ покерных комбинаций, тесты на перекрытие и другие нетривиальные проверки.

Краткое сравнение тестов представлено в таблице 1.

Таблица 1

Сравнение статистических тестов

Тест	Простота реализации	Чувствительность	Применение
Хи-квадрат	Высокая	Средняя	Базовая проверка
NIST SP 800-22	Средняя	Высокая	Криптографическая оценка
DIEHARD	Низкая	Очень высокая	Глубокий анализ

Тест хи-квадрат позволяет формально проверить гипотезу о равномерности распределения битов и выявить скрытые зависимости между элементами последовательности. Его преимущество заключается в строгой математической основе, объективности и возможности автоматизации.

Таким образом, тест хи-квадрат представляет собой состоятельный и практичный инструмент для анализа случайности, который может быть рекомендован для применения на этапах разработки и тестирования криптографических систем.

РЕАЛИЗАЦИЯ И РЕЗУЛЬТАТЫ ЭКСПЕРИМЕНТАЛЬНОГО АНАЛИЗА

Для практической проверки эффективности статистического подхода к оценке криптостойкости была разработана программная реализация теста хи-квадрат, предназначенная для анализа бинарных последовательностей различной длины и структуры.

Описание программной реализации

Программа была написана на языке Python версии 3.10 с использованием следующих библиотек:

- **scipy.stats** — для расчёта значения χ^2 и p-значения;

• **numpy** — для работы с массивами данных;

• **tabulate** — для форматирования вывода в виде таблиц;

• **sklearn.neural_network.MLPClassifier** — для сравнительного анализа.

Анализ проводился по следующей схеме.

1. Битовая последовательность считывается из файла.

2. Последовательность разбивается на окна фиксированного размера: 8, 16 и 32 бита.

3. Для каждого окна рассчитывается частота появления следующего или предыдущего бита.

4. Вычисляется значение теста хи-квадрат и соответствующее p-значение.

5. Результаты группируются и выводятся в табличном виде.

Размер обучающей выборки составлял 20% от доступных данных, но не более 50 битов. Это позволило минимизировать влияние шума при сохранении репрезентативности выборки [14].

Параметры эксперимента

Эксперимент проводился на бинарной последовательности длиной 1140230 бит, которая была протестирована по следующим параметрам:

• **размеры окон:** 8, 16 и 32 бита;

• **направление анализа:** вперёд (следующий бит) и назад (предыдущий бит);

• **количество тестов:** от 25 до 128939 на каждый размер окна;

• **метрики:** среднее значение χ^2, среднее p-значение.

Результаты тестирования

В результате проведённого анализа были получены следующие показатели (отражены в таблице 2).

Таблица 2

Результаты теста хи-квадрат

Размер окна	Направление	Тестов	Среднее χ^2	Среднее p
8	вперед	82777	χ^2=0.48	p=0.501
8	назад	128939	χ^2=0.48	p=0.502
16	вперед	346	χ^2=0.50	p=0.480
16	назад	514	χ^2=0.50	p=0.480
32	вперед	25	χ^2=0.50	p=0.480
32	назад	25	χ^2=0.50	p=0.480

Интерпретация результатов

В рамках примера таблицы 2 результаты можно интерпретировать следующим образом:

Общее среднее χ^2: 0.48

Общее среднее p-значение: 0.501

Вывод по методу хи-квадрат: последовательность демонстрирует признаки криптографической стойкости – P-значение (0.501) больше 0.05.

Анализ значений p-величины показал, что для всех исследуемых размеров окон среднее значение превышает уровень значимости p=0.05, что указывает на высокую степень случайности исследуемой последовательности.

Это позволяет сделать вывод о её потенциальной криптографической стойкости.

Среднее значение теста хи-квадрат также остаётся близким к теоретически ожидаемому ($\chi^2 \approx 0.5$), что подтверждает равномерное распределение частот и отсутствие явных закономерностей в битовой последовательности.

Также важно отметить:

• Разница между направлениями анализа ("вперёд" и "назад") незначительна, что говорит об отсутствии направленной зависимости.

• Увеличение размера окна не привело к снижению уровня случайности, что свидетельствует о хорошем качестве ГПСЧ.

Особенности программной реализации

Важными особенностями являются:

• возможность автоматизации анализа без участия человека;

• поддержка различных размеров окон;

• прозрачность вычислений и возможность воспроизведения результатов;

• интеграция с классическими статистическими метриками.

Подход, основанный на скользящем окне, позволяет повысить чувствительность теста хи-квадрат и обнаружить скрытые зависимости между соседними битами. При этом он отличается простотой реализации и может быть рекомендован как инструмент первичной оценки качества генераторов псевдослучайных чисел.

ПЕРСПЕКТИВЫ РАЗВИТИЯ СТАТИСТИЧЕСКИХ МЕТОДОВ

В условиях постоянного развития криптографии и роста вычислительных возможностей злоумышленников, критически важно совершенствовать существующие подходы к оценке криптостойкости генераторов псевдослучайных чисел. Статистические методы, и, в частности, тест хи-квадрат, остаются важным инструментом анализа случайности, однако они также нуждаются в модификации и адаптации под современные требования.

Интеграция с блочным анализом

Одним из направлений развития является переход от побитового анализа к блочному. В отличие от традиционного подхода, при котором последовательность разбивается на фиксированные окна и проверяется поэлементная случайность, блочный анализ позволяет выявлять скрытые зависимости между группами битов, что особенно важно для ГПСЧ, генерирующих данные фиксированной разрядности (например, 8-, 16-, 32-битные).

Блоковый подход увеличивает чувствительность теста к сложным паттернам, которые не обнаруживаются при обычном применении теста хи-квадрат. Например, вместо проверки частоты появления нулей и единиц, можно анализировать частоту появления определённых комбинаций битов внутри блока. Это даёт возможность:

- определить повторяемость конкретных шаблонов;
- обнаружить закономерности в распределении блоков;
- повысить точность анализа за счёт большего объёма данных на итерацию.

Этот метод был частично реализован в экспериментальной части работы, где проводился анализ бинарной последовательности с использованием окон размером 8, 16 и 32 бита. Полученные результаты продемонстрировали устойчивость равномерного распределения и отсутствие явной корреляции между блоками.

Комбинирование с другими статистическими тестами

Тест хи-квадрат эффективен для проверки равномерности распределения, но не всегда способен обнаруживать более сложные виды неслучайности, такие как временные зависимости или предсказуемость на основе истории последовательности. Для повышения надежности рекомендуется комбинировать его с другими статистическими тестами. Далее представлены примеры тестов.

- Частотный тест NIST – проверяет долю нулей и единиц в последовательности.
- Тест на самые длинные серии в блоке – выявляет аномально длинные цепочки одинаковых битов.
- Тест ранговой корреляции матрицы – позволяет оценить наличие линейных зависимостей между битами.
- Тест на совпадение блоков – анализирует повторяемость фрагментов последовательности.

Комбинирование тестов позволяет создать многоуровневую систему диагностики, в которой каждый тест направлен на выявление специфического типа неслучайности. Такой подход повышает общую достоверность оценки криптостойкости.

Возможности автоматизации и масштабируемости

Еще одной перспективой является автоматизация процесса анализа и масштабируемость методов для работы с длинными последовательностями и потоковыми данными. Благодаря простой алгоритмической структуре тест хи-квадрат легко поддается параллелизации и может быть внедрен в системы реального времени для мониторинга качества выходной последовательности ГПСЧ.

Кроме того, возможно развитие адаптивных стратегий, при которых параметры тестирования (размер окна, уровень значимости, частота проверок) изменяются динамически в зависимости от характеристик исследуемой последовательности.

Сравнение с машинным обучением

Хотя данная работа сосредоточена именно на статистическом подходе, стоит отметить, что в ряде исследований рассматривается использование методов машинного обучения для оценки криптостойкости. Несмотря на их высокую чувствительность, эти методы требуют значительных вычислительных ресурсов и могут давать менее интерпретируемые результаты. Поэтому, особенно на этапах первичного анализа и сертификации, целесообразнее использовать хорошо формализованные и воспроизводимые статистические тесты.

Однако для будущих исследований представляет интерес гибридный подход, при котором нейросетевые модели используются для предварительного анализа, а статистические тесты — для верификации и формальной проверки гипотез.

Таким образом, статистические методы, и в первую очередь тест хи-квадрат, остаются актуальными и состоятельными средствами оценки криптостойкости ГПСЧ. Их развитие связано с переходом к блочному представлению данных, комбинированием с другими тестами и автоматизацией процесса анализа. Эти усовершенствования позволяют сохранить преимущества классической статистики — объективность и воспроизводимость — при этом повышая эффективность и применимость к современным задачам криптографии.

ЗАКЛЮЧЕНИЕ

В рамках проведенного исследования были рассмотрены и экспериментально проверены статистические методы оценки криптостойкости генераторов псевдослучайных чисел (ГПСЧ), с акцентом на тест хи-квадрат как один из наиболее универсальных и математически обоснованных инструментов анализа случайности.

Тест хи-квадрат показал свою состоятельность при проверке гипотезы о равномерности распределения битовых последовательностей. Анализ проводился по окнам фиксированного размера: 8, 16 и 32 бита, что позволило повысить чувствительность к зависимостям между соседними элементами. Полученные значения среднего χ^2=0.48–0.50 и p=0.480–0.513 для всех исследуемых размеров окон свидетельствуют о соответствии исследуемой последовательности гипотезе случайности. Это позволяет рекомендовать её для использования в криптографических системах.

Подход, основанный на скользящем окне, обеспечивает объективную и воспроизводимую оценку случайности и может быть внедрён в системы автоматизированного тестирования ГПСЧ. Простота реализации и возможность масштабирования делают его особенно актуальным на этапах первичной верификации качества генераторов псевдослучайных чисел.

Результаты экспериментального анализа подтверждают эффективность предложенного подхода и указывают на потенциал применения статистических методов для формализованной оценки криптостойкости. В отличие от сложных нейросетевых моделей, тест хи-квадрат не требует значительных вычислительных ресурсов и позволяет получить интерпретируемый результат, что особенно важно при сертификации криптографических алгоритмов.

Особое внимание уделено программной реализации, позволяющей автоматизировать процесс анализа и использовать разные размеры окон для повышения точности. Такая гибкость параметров открывает возможности для адаптации метода под различные типы ГПСЧ и задачи информационной безопасности.

Таким образом, статистические методы, и, в частности, тест хи-квадрат, остаются актуальными и состоятельными средствами оценки криптостойкости ГПСЧ. Они обеспечивают формализованную, объективную и легко автоматизируемую оценку случайности, что делает их подходящими для широкого применения в криптографии и информационной безопасности. Предложенный подход может быть рекомендован к использованию в качестве основы для стандартизованных методик тестирования генераторов псевдослучайных чисел.

СПИСОК ЛИТЕРАТУРЫ

1. Мирзоян С. А. Применение нейронных сетей для оценки криптостойкости генераторов псевдослучайных чисел // Hi-Hume Journal. 2025. №1(8). С. 24–31.
2. Stallings W. Cryptography and Network Security: Principles and Practice. Pearson, 2017. 766 p.
3. Menezes A., van Oorschot P., Vanstone S. Handbook of Applied Cryptography. CRC Press, 1996. 795 p.
4. Knuth D. E. The Art of Computer Programming, Volume 2: Seminumerical Algorithms. Addison-Wesley, 1997. 782 p.
5. Rukhin A., Soto J., Nechvatal J. et al. A Statistical Test Suite for Random and Pseudorandom Number Generators for Cryptographic Applications. NIST Special Publication 800–22 Revision 1a, 2010. 131 p.
6. Anderson R. Security Engineering: A Guide to Building Dependable Distributed Systems. Wiley, 2020. 1232 p.
7. Matsumoto M., Nishimura T. Mersenne Twister: A 623-dimensionally equidistributed uniform pseudo-random number generator // ACM Transactions on Modeling and Computer Simulation. 1998. Vol.8, No.1. Pp.3–30.
8. Ferguson N., Schneier B., Kohno T. Cryptography Engineering: Design Principles and Practical Applications. Wiley, 2010. 385 p.
9. L'Ecuyer P., Simard R. TestU01: A C Library for Empirical Testing of Random Number Generators // ACM Transactions on Mathematical Software. 2007. Vol.33, No.4. Article 22. 40 p.

10. Мирзоян С. А. Применение технологий искусственного интеллекта для повышения качества генераторов псевдослучайных чисел // Вестник современных цифровых технологий. 2025. №22. С. 38–44.
11. National Institute of Standards and Technology. Digital Signature Standard (DSS) // FIPS PUB 186-4, 2013. 131 p.
12. Marsaglia G. Diehard: a battery of tests of randomness. Florida State University, 1996. URL: https://cir.nii.ac.jp/crid/1571698600935841152 (Дата обращения: 17.07.2025).
13. Pedregosa F. et al. Scikit-learn: Machine Learning in Python // Journal of Machine Learning Research. 2011. Vol.12. Pp.2825–2830.
14. McKinney W. Data Structures for Statistical Computing in Python // Proceedings of the 9th Python in Science Conference. 2010. Pp.51–56.

УДК: 004.056

Математическая модель для оценки уровня зрелости системы кибербезопасности организации

E.S. Polikarpov, D.V. Lipendin

Mathematical Model for Assessing the Maturity Level of an Organization's Cybersecurity System

Abstract. *The article considers the issues of determining the maturity of an organization's cybersecurity based on indicators related to available resources, external conditions and internal factors. A mathematical model is provided for quantitative assessment of the maturity of an organization's information security. The dependence of the maturity level on resources, internal and external factors, the state of maturity of the cybersecurity system, as well as the level of training of personnel of departments responsible for this area is obtained.*

Keywords: *cybersecurity, cybersecurity maturity level, cybersecurity maturity assessment, mathematical modeling.*

Е.С. Поликарпов[1]
Д.В. Липендин[2]
[1]Кандидат технических наук, доцент кафедры международной информационной безопасности, Московский государственный лингвистический университет.
E-Mail: binox@mail.ru
[2]Аспирант кафедры международной информационной безопасности, Московский государственный лингвистический университет.
E-Mail: lipendindv@gmail.com

Аннотация. *В статье рассматриваются вопросы определения зрелости кибербезопасности организации на основе показателей, связанных с доступными ресурсами, внешними условиями и внутренними факторами. Приводится математическая модель для количественной оценки зрелости информационной безопасности организации. Получена зависимость уровня зрелости от ресурсов, внутренних и внешних факторов, текущего уровня зрелости системы кибербезопасности, а также уровня подготовки персонала подразделений, ответственных за данное направление.*

Ключевые слова: *кибербезопасность, уровень зрелости кибербезопасности, оценка зрелости кибербезопасности, математическое моделирование.*

ВВЕДЕНИЕ

В современном цифровом мире кибербезопасность становится одним из ключевых факторов устойчивого развития и конкурентоспособности организаций. С ростом числа и сложности киберугроз, а также с увеличением объема обрабатываемых данных, компании сталкиваются с необходимостью не только внедрять базовые меры защиты, но и системно развивать свои возможности в области информационной безопасности. Поскольку поверхности атак организаций растут, а ландшафт угроз становится все более сложным, важно, чтобы организации не только имели установленную программу развития кибербезопасности, но и способ оценки и наращивания ее зрелости с течением времени.

Зрелая система кибербезопасности может способствовать снижению риска нарушений информационной безопасности в организации, защите конфиденциальной информации и обеспечению соответствия юридическим и нормативным требованиям.

Уровень зрелости кибербезопасности отражает степень готовности организации эффективно противостоять угрозам, управлять рисками и обеспечивать защиту критически важных активов. Введение модели зрелости позволяет оценить текущий статус безопасности, выявить слабые места и определить направления для дальнейшего совершенствования.

В этом контексте создаваемые модели зрелости служат инструментами оценки возможностей улучшения состояния информационной безопасности для конкретной организации. Основными компонентами типовой модели оценки зрелости являются средства для оценки и сравнительного анализа производительности, дорожная карта (план работ) для улучшения состояния кибербезопасности, а также средства выявления пробелов в организации безопасности и разработки планов улучшения состояния кибербезопасности.

Анализ зрелости кибербезопасности может проводиться на основе разных алгоритмов, и полученная оценка может принимать разные формы в зависимости от используемой модели. Например, в модели CMMC (от англ. Cybersecurity Maturity Model Certification – сертификация модели зрелости кибербезопасности) в результате оценки организации присваивается один из трех уровней зрелости на основе определенных характеристик, описанных

в том числе в стандартах NIST. Модель используется подрядчиками в оборонной промышленной базе для соответствия требованиям безопасности из NIST SP 800-171.

Другой пример – модель O-ISM3, оценивает зрелость функционирования существующих процессов системы управления информационной безопасностью (СУИБ) организации. Эти процессы классифицируются по пяти уровням зрелости в соответствии с определенными метриками. O-ISM3 (аббр. от Open Information Security Management Maturity Model) направлена на обеспечение процессов безопасности организации для того, чтобы они были реализованы на уровне, соответствующем бизнес-требованиям.

Существуют и другие подходы и модели, имеющие свои достоинства и недостатки, например FedRAMP, FIPS, CMMI и т.д. Большинство таких подходов основываются на мнении экспертов в конкретной области.

В настоящей работе рассматривается базовая математическая модель для оценки зрелости кибербезопасности организации. Исследуемая модель учитывает скорость изменения зрелости организации, которая зависит от ресурсов, внутренних факторов, внешних факторов и текущего уровня зрелости, а также уровня подготовки персонала структурных подразделений. Результаты математического моделирования проанализированы с использованием среды Mathcad, определена зависимость зрелости организации от времени при влиянии различных факторов и условий.

ОБ ОСНОВНЫХ ПОЛОЖЕНИЯХ ОЦЕНКИ ЗРЕЛОСТИ КИБЕРБЕЗОПАСНОСТИ

Уровень зрелости любого процесса (также встречается термин «уровень возможностей бизнес-процесса») чаще всего определяют, как показатель процесса, характеризующий его способность соответствовать текущим и будущим целям предприятия. Зрелость показывает, насколько процесс управляем и прогнозируем [1].

Существует разделение зрелости на цифровую зрелость и зрелость кибербезопасности. Категория цифровой зрелости позволяет формулировать выводы о динамике, формах и векторах развития цифровой экономики [2].

Само понятие «зрелость» в контексте цифровой экономики может означать период времени и некое состояние, в котором находится участник социальных и/или экономических отношений и которое позволяет говорить о том, что достижение поставленных субъектом целей в полной мере осуществляется с помощью технологий цифровой экономики. [2]

При оценке зрелости информационной безопасности целесообразна комбинация определенных оценочных методов, где есть место мерам оценки правильности (например, оценки соответствия требованиям) и эффективности (например, измерения степени возможности самосовершенствования процессов и их развития) [3].

Если рассматривать цифровую зрелость, то ее также связывают с процессом цифровой трансформации [4].

Модели оценки зрелости встречаются в стандартах и оказываются необходимыми в области безопасности. Существуют исследования, направленные на воспроизведение и адаптацию концепции зрелости к информационной безопасности или безопасности ИКТ. Это побудило организации к созданию множества моделей оценки зрелости кибербезопасности, значительно расширив их выбор [5].

Модели оценки зрелости могут применяться к различным задачам кибербезопасности, например, для противодействия инсайдерским угрозам информационной безопасности [6].

Разработано множество моделей зрелости, которые применяются именно к информационной безопасности [7]. В моделях зрелости информационной безопасности применяются разные подходы и рассматриваются разные области деятельности организации, например, корпоративное управление, культура организации, управление сервисом, архитектура системы [8].

Модели зрелости информационной безопасности могут опираться на разные системы стандартов. Например, специальная модель зрелости информационной безопасности разработана для системы кибербезопасности NIST [9]. В других случаях модель зрелости информационной безопасности может опираться на стандарты серии ISO27K для выделения определенных областей, которые будут рассматриваться [10].

В данной статье речь идет об обобщенном понятии зрелости кибербезопасности.

ПОСТРОЕНИЕ УПРОЩЕННОЙ МОДЕЛИ ОЦЕНКИ ЗРЕЛОСТИ В MATHCAD

Модели зрелости кибербезопасности — это структуры, которые помогают организациям сравнивать свои текущие возможности в области безопасности, а также цели и приоритеты идентификации для продвижения к более высоким уровням

зрелости. Современные модели должны учитывать максимальное количество условий и факторов, действие которых меняется с течением времени.

Чем точнее описать условия и факторы, влияющие на зрелость, тем точнее можно определить уровень зрелости. В предлагаемой модели присутствуют следующие элементы:

M(t) – зрелость кибербезопасности;

R(t) – доступные ресурсы (например, финансирование, персонал, технологии);

E(t) – внешние условия (рыночная среда, конкуренция);

I(t) – внутренние факторы (уровень знаний, эффективность процессов).

Будущие исследования покажут, как измерять или оценивать эти параметры с течением времени. Используемые технологии могут устаревать со временем, что должно в свою очередь снижать уровень зрелости и создавать угрозы информационной безопасности. Также необходимо периодически повышать и обновлять компетенцию работников структурного подразделения.

Для оценки изменения уровня зрелости кибербезопасности используем дифференциальное уравнение следующего вида:

$$\frac{dM}{dt} = k_1 R(t) + k_2 E(t) + k_3 I(t) - k_4 M(t)$$

Где dM/dt – скорость изменения зрелости кибербезопасности.

Коэффициенты k_1, k_2, k_3, определяющие вклад ресурсов, внешних условий и внутренних факторов соответственно.

Коэффициент k_4 – коэффициент затухания, который учитывает возможное снижение зрелости из-за устаревания технологий и других факторов.

Дальнейшие исследования помогут ответить на вопросы, как определить эти коэффициенты, дают ли они в сумме единицу и могут ли быть связаны внутренние факторы и слагаемое $k_4M(t)$. Например, устаревшее оборудование также относится к внутренним факторам.

В упрощенной модели, рассматриваемой нами в настоящей работе, ресурсы и внешние условия постоянны ($R(t)=R_0$, $E(t)=E_0$) и внутренние факторы зависят от текущего уровня зрелости ($I(t)=k_5M(t)$). Уравнение принимает вид:

$$\frac{dM}{dt} = k_1 R_0 + k_2 E_0 + k_3 k_5 M(t) - k_4 M(t)$$

Полученная упрощенная модель была проанализирована в среде Mathcad. Так как наша модель основана на дифференциальном уравнении, для решения были использованы блоки решения и функции Odesolve. Для создания блока решения используется ключевое слово Given.

Далее мы перечислили все коэффициенты, так как они являются постоянными. Значения для этих коэффициентов можно получить, например, с помощью экспертной оценки того, насколько те или иные факторы влияют на скорость изменения зрелости кибербезопасности. Эксперты оценивают это влияние по шкале от нуля до единицы.

Предположим, что в ходе экспертной оценки были получены случайные коэффициенты, представленные в диапазоне от 0 до 1. Затем вводится дифференциальное уравнение из упрощенной модели. После этого параметрам R_0 и E_0 присваивается значение 1. Результаты дальнейших исследований могут содержать потенциальные способы измерения этих параметров.

Далее задаются начальные условия для решения дифференциального уравнения (чему равны M(0) и M’(0)). Дополнительное исследование может показать, как получить начальные условия. В данном примере M(0) = 0, то есть зрелость в нулевой момент времени равна нулю. После этого используется функция Odesolve. Ниже на рисунке 1 приведен полученный график функции зрелости M(t).

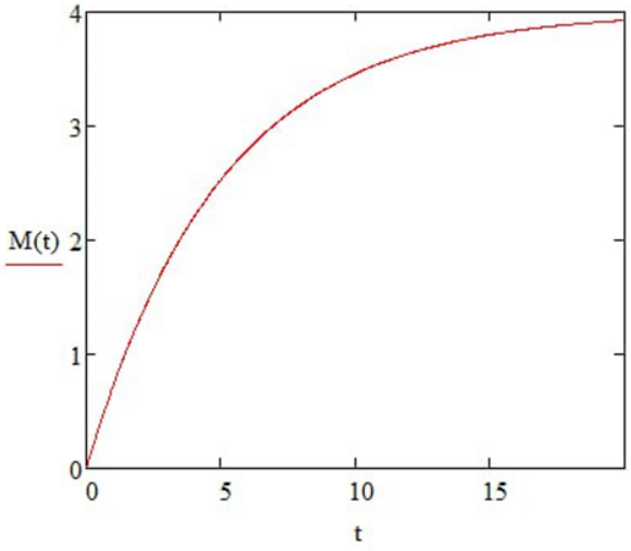


Рис. 1. Зависимость зрелости кибербезопасности организации от времени при влиянии различных факторов и условий

Опираясь на решение дифференциального уравнения и интерпретацию этого решения, мы получаем, что при a>0 ($a=k_4-k_3k_5$, $b=k_1R_0+k_2E_0$) зрелость M(t) стремится к значению b/a, то есть в данном примере – к значению 4. На графике видно, что с течением времени зрелость растет и стремится к значению 4.

ПРИМЕР РАСЧЕТА ПО ПРЕДЛАГАЕМОЙ МОДЕЛИ

Чтобы провести расчет по предлагаемой модели, необходимо определить значения для коэффициентов k_1, k_2, k_3, k_4, k_5 и параметров R_0 и E_0. Предположим, что значения для коэффициентов были получены с помощью экспертной оценки. В ходе такой оценки эксперты ответили на вопрос, насколько те или иные факторы влияют на скорость изменения зрелости кибербезопасности. Эксперты оценили это влияние по шкале от нуля до единицы: $k_1 = 0.5$, $k_2 = 0.3$, $k_3 = 0.8$, $k_4 = 0.6$, $k_5 = 0.5$. Далее предположим, что мы провели анализ ресурсов организации и внешних условий организации. В ходе этого анализа мы определили значения для параметров R_0 и E_0. $R_0 = 150$ и $E_0 = 20$. Зная эти значения, мы можем провести расчет a и b ($a=k_4-k_3k_5$, $b=k_1R_0+k_2E_0$).

$a = 0.6 – 0.8 * 0.5 = 0.2$.

$b = 0.5 * 150 + 0.3 * 20 = 81$.

Значение a>0, значит зрелость M(t) стремится к значению b/a.

$b/a = 81/0.2 = 405$.

Если внести исходные данные в среду Mathcad, можно получить график функции зрелости M(t). На рисунке 2 приведен полученный график.

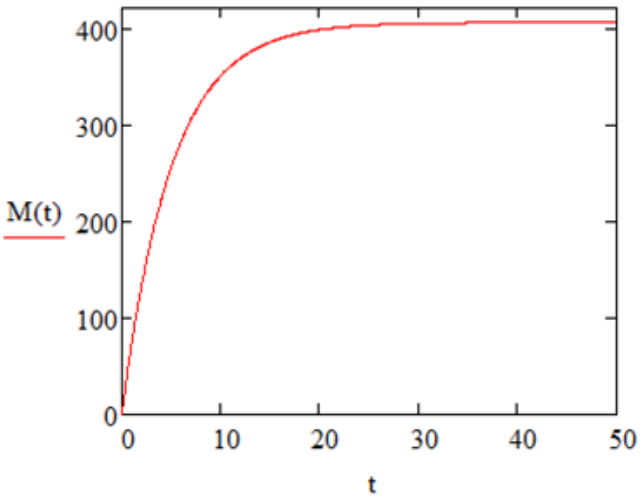


Рис. 2. Пример зависимости зрелости кибербезопасности организации от времени при влиянии различных факторов и условий

ПРЕИМУЩЕСТВА ПОСТРОЕННОЙ МОДЕЛИ

Предлагаемая упрощенная модель оценки зрелости кибербезопасности организации имеет следующие преимущества:

1. Использование числовых параметров позволяет получать количественную оценку уровня зрелости, что повышает точность и объективность анализа по сравнению с традиционными качественными методами.

2. Внедрение дифференциального уравнения в модель обеспечивает динамическое описание процесса развития и изменения уровня кибербезопасности во времени, учитывая его непрерывную природу и возможные колебания.

3. В результате применения модели пользователь получает график функции зрелости, что является не просто статической оценкой, а динамической картиной развития зрелости кибербезопасности. Это способствует более точному планированию и принятию управленческих решений.

ЗАКЛЮЧЕНИЕ

В статье представлена математическая модель оценки зрелости кибербезопасности с учетом таких параметров, как ресурсы организации, внешние и внутренние факторы, коэффициент затухания. Данная модель в виде дифференциального уравнения была проанализирована в Mathcad.

Таким образом, получен новый отечественный метод для количественной оценки зрелости кибербезопасности. Его внедрение позволит предприятиям более точно измерять степень собственной готовности к киберугрозам, своевременно выявлять пробелы в информационной инфраструктуре, корректировать планы развития и совершенствования системы управления информационной безопасностью. Оценку зрелости кибербезопасности можно осуществлять на постоянной основе в целях контроля реализации плана развития системы кибербезопасности организации.

Наиболее полезным преимуществом использования модели является ее способность динамического отслеживания состояния зрелости. Кроме того, описанная модель довольно гибкая для всех процессов и систем. Такая модель оценки зрелости кибербезопасности сможет послужить твердой основой для оценки и повышения зрелости отдельных процессов обеспечения информационной безопасности в организации.

Внедряя данную модель, организации могут добиться значительных результатов, включая улучшение согласованности процессов, снижение эксплуатационных расходов на информационную инфраструктуру, повышение эффективности управления информационной безопасностью.

СПИСОК ЛИТЕРАТУРЫ

1. Дмитриева М. А. Применение анализа зрелости информационной безопасности в системе оценки зрелости бизнес-процессов предприятия в целом // Информационная безопасность регионов.- 2015.- №3 (20). – С. 20-24.

2. Погорельцев А. С., Салимьянова И. Г. Особенности оценки цифровой зрелости организаций // Известия СПбГЭУ.- 2022.- №5-2 (137). – С. 118-125.

3. Голованов В. Б. Модель зрелости как подход измерения эффективности процессов информационной безопасности / В. Б. Голованов // Труды Международного симпозиума «Надежность и качество». – 2006.- № 2.- С. 159-161.

4. Попов Е. В., Симонова В. Л., Черепанов В. В. Уровни цифровой зрелости промышленного предприятия // Journal of new economy.- 2021.- №22-2. – С. 88-109.

5. Rabii, Anass, et al. Information and cyber security maturity models: a systematic literature review // Information & Computer Security. 2020, Vol.28. Iss.4. Pp.627-644.

6. Поляничко М. А. Применение модели зрелости для противодействия инсайдерским угрозам информационной безопасности // МНИЖ.- 2019.- №4-1 (82). – С. 57-59.

7. Милославская Н. Г., Сагиров Р.А. Обзор моделей зрелости процессов управления информационной безопасностью // Безопасность информационных технологий. – 2015. – Том 22, № 2. – С. 76-84.

8. Saleh Malik F. Information security maturity model // International Journal of Computer Science and Security (IJCSS) 2011, Vol.5. Iss.3. P.21.

9. Almuhammadi Sultan, Majeed Alsaleh. Information security maturity model for NIST cyber security framework // Computer Science & Information Technology (CS & IT). 2027. Vol.7. Iss.3 Pp.51-62.

10. Spruit M., Röling M. ISFAM: The Information Security Focus Area Maturity Model / ResearshGate. 2014, January. URL: https://www.researchgate.net/publication/288134391_ISFAM_The_information_security_focus_area_maturity_model (Дата обращения: 06.09.2025)

УДК: 51, 003.26

Симметрия в криптографии

I. A. Kirillov

Symmetry in cryptography

Abstract. *The article discusses the general concept and variants of definitions of symmetry in specific areas of science. The general definition of the concept treats symmetry primarily as a philosophical category and practically does not reflect the essence of the phenomenon from the standpoint of mathematics. The issues of symmetry in cryptography are considered. Certain inaccuracies in the definition and use of the concept of symmetry in some cryptographic publications are established.*

Keywords: *symmetry, cryptography, symmetric encryption, invariance.*

И.А. Кириллов
Кандидат технических наук, заслуженный профессор МГЛУ, доцент кафедры международной информационной безопасности Института информационных наук, Московский государственный лингвистический университет.
E-mail: I.A.Kirillov@gmail.com

Аннотация. *В статье обсуждается общее понятие и варианты определений симметрии в конкретных областях науки. Общее определение понятия трактует симметрию прежде всего как философскую категорию и практически не отражает сути явления с позиции математики. Рассматриваются вопросы симметрии в криптографии. Устанавливаются отдельные неточности в определении и использовании понятия симметрии в некоторых публикациях в области криптографии.*

Ключевые слова: *симметрия, криптография, симметричное шифрование, инвариантность.*

ОБЩЕЕ ПОНЯТИЕ СИММЕТРИИ

«Несмотря на большую литературу о симметрии и на огромные практические приложения, очень нелегко выяснить положение симметрии в системе наук. О ней говорят, как о чем-то общеизвестном, самопонятном и делают из нее выводы, которыми пользуются на каждом шагу. Но мы не найдем в этой литературе точного определенного указания на то, что же представляют собой явления симметрии в природных процессах…» [1]. Этому высказыванию выдающегося российского ученого Владимира Ивановича Вернадского уже более восьмидесяти лет, но разночтения, связанные с понятием симметрии, и сегодня нередки, в том числе и в области криптографии.

Приведем определение из Википедии: «Симметрия (др.-греч. Συμμετρία = «соразмерность»; от συν- «совместно» + μετρέω «мерю»), в широком смысле — соответствие, неизменность (инвариантность), проявляемые при каких-либо изменениях, преобразованиях (например: положения, энергии, информации, другого)[1]».

Попытка уточнения этого весьма краткого и малоконкретного понятия наталкивается на необходимость указания конкретной области, с которой симметрия связана, будь то физика, химия, биология, кристаллография, криптография и т.д.

ПОНЯТИЕ СИММЕТРИИ В МАТЕМАТИЧЕСКОМ АСПЕКТЕ

Приведем пример определения с учетом особенностей математики (точнее, геометрии на плоскости) из Большой российской энциклопедии: «Симметрия в математике (лат. symmetria, греч. συμμετρία – соразмерность):

1) Симметрия на плоскости (в узком смысле), или (зеркальное) отражение относительно прямой α – преобразование … плоскости, при котором каждая точка M переходит в точку M′ такую, что отрезок MM′ перпендикулярен прямой α и делится ею пополам. Прямая α называется осью симметрии.

2) Симметрия на плоскости (в широком смысле) – свойство геометрической фигуры Ф совмещаться с собой при действии некоторой группы G. В данном случае G - это множество различных преобразований, обладающих определенными свойствами, называемое группой симметрии Ф.

При этом преобразования … сохраняют длины векторов и углы между ними. Таким образом, симметрия отражает некоторую правильность формы фигуры, её инвариантность при действии преобразований из G. Например, если фигура Ф на плоскости такова, что повороты относительно некоторой точки O на угол 360°/n, (где n≥2 – целое число), переводят её в себя, то говорят, что Ф обладает

[1] Симметрия. URL: https://ru.wikipedia.org/wiki/Симметрия (Дата обращения 25.05.2025).

симметрией n-го порядка, а O называют центром симметрии n-го порядка. … Окружность обладает симметрией бесконечного порядка, поскольку совмещается с собой при повороте вокруг центра на любой угол.

Простейшими видами симметрии на плоскости, помимо симметрий, порождённых отражениями и поворотами … являются … симметрии переноса; в этом случае фигура совмещается с собой переносом вдоль некоторой прямой (оси переноса) на некоторый отрезок. Фигура с одной осью переноса обладает бесконечным множеством плоскостей симметрии, перпендикулярных оси переноса, поскольку любой перенос можно осуществить двумя последовательными отражениями.

Комбинации симметрий, порождённые отражениями и вращениями (исчерпывающие простейшие виды симметрий конечных фигур), а также переносами, представляют интерес и являются предметом исследования в различных областях естествознания, искусства и т. д[2]»

В добавление к приведенным выше определениям хотелось бы представить ещё и определение советского и российского ученого Ю.А.Урманцева[3]: «Симметрия — это категория, обозначающая признаки «П» («инварианты») объектов (системы) „О“ вместе с такими преобразованиями „И“, которые объекты „О“ по признакам «П» оставляют тождественными самим себе» [2]. Точным математическим выражением симметрии является особая алгебраическая структура — группа преобразований (множество автоморфных преобразований с внутренней бинарной ассоциативной операцией над ними).

Из этого определения усматривается непосредственная связь между понятиями «симметрия» и «инвариантность», которая в виде теоремы была доказана Эмми Нётер в 1918 году. Этой теоремой устанавливалось, что «каждой непрерывной симметрии физической системы соответствует некоторый закон сохранения (инвариантности)»[4].

СИММЕТРИЯ В КРИПТОГРАФИИ

Обратимся теперь к симметрии в криптографии. «Симметричные криптосистемы (также симметричное шифрование, симметричные шифры) — способ шифрования, в котором для шифрования и расшифрования применяется один и тот же криптографический ключ. До изобретения схемы асимметричного шифрования единственным существовавшим способом являлось симметричное шифрование»[5].

Аналогичное определение приводится во многих изданиях по криптографии, в частности в книге [3, с.9]: «В симметричных криптосистемах для зашифрования и расшифрования используется один и тот же ключ».

Менее категоричное, и как будет подтверждено далее, более корректное определение приводится в учебном пособии [4], высокое качество которого ко всему прочему подтверждено грифом «Допущено Министерством образования Российской Федерации в качестве учебного пособия для студентов высших учебных заведений, обучающихся по группе специальностей в области информационной безопасности».

На стр.59 указанного пособия читаем: «Различают симметричные и асимметричные криптосистемы. В симметричных системах знание ключа зашифрования k_1 позволяет легко найти ключ расшифрования k_2 (в большинстве случаев эти ключи просто совпадают)» [4, с.59].

Продолжая констатацию дефиниций вслед за авторами [4, 5] воспользуемся следующей записью алгебраической модели шифра (шифрсистемы):

$S = (X,K,Y,E,D)$.

Приведем пояснения используемых обозначений [4, с.75]:

$X = \{x\}$ – множество возможных открытых текстов;

$K = \{k\}$ – множество возможных ключей;

Y – множество соответствующих шифрованных текстов;

E – множество преобразований зашифрования;

D – множество преобразований расшифрования;

$E= \{E_k: k \in K\}$, $E_k: K \times X \to Y$;

$D= \{D_k: k \in K\}$, $D_k: K \times Y \to X$;

$k = (k_e,k_d)$, k_e – ключ зашифрования, k_d – ключ расшифрования;

$E_k \equiv E_{ke}$, $D_k \equiv D_{kd}$;

Для любых $k \in K$ и $x \in X$ выполняется равенство $D_k(E_k(x))=x$;

Для любых $y \in Y$ существуют такие $x \in X$ и $k \in K$, что $y=E_k(x)$.

Без потери общности можно считать, что как открытые, так и шифрованные тексты представляют собой последовательности символов некоторого множества A ($|A| = n$), называемого алфавитом.

Попытаемся наполнить общее определение симметрии, данное Урманцевым Ю. А., конкретным криптографическим содержанием, приведенным в [4, с. 75]. При этом естественно отождествить объект

[2] Симметрия. Большая российская энциклопедия. URL: https://bigenc.ru/c/simmetriia-26c9e1?ysclid=mawa2nvgmt657084738 (Дата обращения 25.05.2025).
[3] Урманцев, Юнир Абдуллович-Википедия. URL: https://ru.wikipedia.org/wiki/Урманцев,_Юнир_Абдуллович (Дата обращения 25.05.2025).
[4] Симметрия (физика) – Википедия. URL: https://ru.wikipedia.org/wiki/Симметрия_(физика) (Дата обращения: 25.05.2025).
[5] Симметричные криптосистемы – Википедия. URL: https://ru.wikipedia.org/wiki/Симметричные_криптосистемы (Дата обращения: 25.05.2025).

(систему) „О“ с криптосистемой О ≈ S = (X,K,Y,E,D). Преобразования „И“, связанные с объектом „О“, включают преобразования E и D объекта S так, что И ≈ {E, D}, но при этих преобразованиях должны оставаться неизменными (инвариантными) признаки П ≈ {$k_e = k_d$}. Таким образом, может быть введено определение симметрии шифрсистемы S относительно инвариантных ключей зашифрования и расшифрования.

Приведенная выше алгебраическая модель шифра может быть использована для конкретизации описания шифров простой замены. Для этого понадобится ввести следующие уточнения для множества возможных ключей K:

$K \subseteq S_n(A)$, где n – мощность алфавита A, $S_n(A)$ — симметрическая группа подстановок множества A [4, с77], при этом каждый элемент $k \in K$ трактуется как некоторая подстановка из $S_n(A)$ или биективное отображение, определяемое этой подстановкой. Если ключ зашифрования $k_e = k$, то ключ расшифрования шифра простой замены $k_d = k^{-1}$.

Открытый и шифрованный тексты шифра простой замены длины L могут быть записаны в виде последовательностей символов x_i, y_i (i=1,L) алфавита A. Таким образом, для открытого текста будет использована запись $x = (x_1, x_2, \ldots x_L)$, а для соответствующего шифрованного текста – запись $y = (y_1, y_2, \ldots y_L)$. При этом правила зашифрования и расшифрования шифра простой замены в алфавите A примут вид:

$E_k(x) = (k(x_1), \ldots, k(x_L))$,

$D_k(y) = (k^{-1}(y_1), \ldots, k^{-1}(y_L))$.

Для еще большей определенности приведенных теоретических выкладок рассмотрим историческую простую замену «Шифр Цезаря» [4, с11]. Гай Светоний Транквилл – биограф Юлия Цезаря – описывал [6] составное кольцо Цезаря, по окружностям двух концентрических частей которого в алфавитном порядке равномерно располагались все буквы латинского алфавита. Выписанные друг под другом, эти алфавиты определяли подстановку k конкретного ключа зашифрования шифра простой замены.

Сдвиг одной концентрической части кольца (соответствующей шифробозначениям) относительно другой его части на три буквенные позиции приводил к конкретной замене, описанной Светонием. Такой вид замены часто называется шифром сдвига Цезаря или просто шифром Цезаря. Хотя в [6] упоминался только шифр Цезаря со сдвигом на три позиции, ясно, что аналогичным образом могут быть реализованы 25 различных шифров, соответствующих 25 различным нетождественным сдвигам.

Следует заметить, что усматриваемая в описанной истории аналогия с упомянутой выше поворотной симметрией 26-го порядка имеет отношение к геометрической форме кольца с алфавитом, но никак не к криптографической симметрии.

Подстановку, соответствующую ключу зашифрования со сдвигом на три для удобства записи представим в виде произведения независимых циклов (в данном случае – в виде одного цикла длины 26) [7]:

k_e = (a, d, g, j, m, p, s, v, y, b, e, h, k, n, q, t, w, z, c, f, i, l, o, r, u, x).

Тогда из определения симметричных криптосистем, приведенного в Википедии должно следовать, что:

k_d = (a, d, g, j, m, p, s, v, y, b, e, h, k, n, q, t, w, z, c, f, i, l, o, r, u, x),

а так как для шифров простой замены должно выполняться соотношение $k_d^{-1} = k_e$, произведение подстановок зашифрования и расшифрования шифра Цезаря должно приводить к тождественной подстановке, что противоречит реально полученному результату:

$k_e k_d$ = (a, g, m, s, y, e, k, q, w, c, i, o, u) (d, j, p, v, b, h, n, t, z, f, l, r, x).

Таким образом, и для общего случая шифров простой замены, к которым относился рассматриваемый пример, нельзя утверждать, что ключи зашифрования и расшифрования совпадают. Следовательно, упомянутое выше определение из Википедии и [3, с.9], приведшее к продемонстрированному противоречию – некорректно. Путем несложных рассуждений, основанных на свойствах симметрических групп, можно установить, что совпадение подстановок зашифрования и расшифрования имеет место только для подстановок, цикловая структура которых не содержит циклы длины более двух.

Для опровержения справедливости приведенного тезиса из Википедии в отношении шифров перестановки: «до изобретения схемы асимметричного шифрования единственным существовавшим способом являлось симметричное шифрование» также может быть построен контрпример. Это следует из необходимости рассмотрения тех же закономерностей симметрической группы ключей шифров перестановки S_L, где L – длина открытого текста.

На первый взгляд рассмотренный вопрос о корректности различных определений симметричности шифров может показаться несущественным. Велика ли разница в том, что «ключи зашифрования и расшифрования совпадают (определение Википедии), или ключи не совпадают, но при известном одном ключе легко отыскивается другой ключ (определение [4, с. 59])»? Однако, как только рас-

сматриваются утверждения о криптографических «ключах», мы должны учитывать основные положения принципа Керкгоффса [4, с.171].

Как известно, «криптографические закономерности» в ряде случаев несопоставимо страшнее «криптографических случайностей». Кроме того, следует привести и последний аргумент: «в серьезных делах мелочей не бывает».

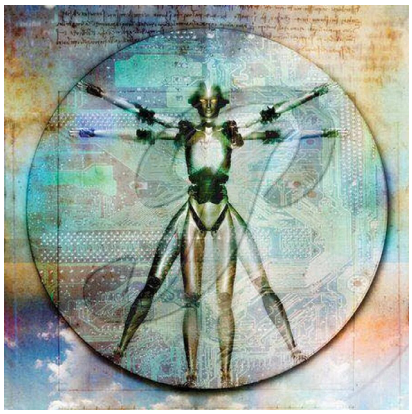
В заключение хотелось бы выразить искреннюю благодарность глубокоуважаемому Юрию Алексеевичу Акулинину за профессиональные и доброжелательные советы и обсуждения.

СПИСОК ЛИТЕРАТУРЫ

1. Вернадский В. И. Философские книги натуралиста (часть 1, раздел: Проблема Времени, Пространства и Симметрии. 1920-1942). М.: Наука, 1988. – 522 с.
2. Урманцев Ю.А. Симметрия природы и природа симметрии: Философские и естественно-научные аспекты. М.: Либроком, 2017. – 232 с.
3. Баричев С. Г., Гончаров В. В., Серов Р. Е. Основы современной криптографии. М.: Горячая Линия - Телеком, 2011. – 175 с.
4. Алферов А.П., Зубов А.Ю., Кузьмин А.С., Черемушкин А.В. Основы криптографии. М.: Гелиос АРВ, 2005. – 480 с.
5. Лось А. Б., Нестеренко А. Ю., Рожков М. И. Криптографические методы защиты информации. Москва: Издательство Юрайт, 2019. – 473 с.
6. Транквилл Г.С. Жизнь двенадцати цезарей. М. Наука 1964г. – 375 с.
7. Глухов М. М., Елизаров В. П., Нечаев А. А. Алгебра: Учебник. — 2-е изд., исправленное и дополненное — СПб.: Издательство «Лань», 2015. – 608 с.

Дневник доцента Ковалёва

Егор Федоров
Писатель, сценарист, драматург
Республика Беларусь



Здравствуйте. Меня зовут Алексей Ковалёв, я доцент кафедры нейробиологии Новосибирского НИИ когнитивных технологий.

Десять лет назад я начал эксперименты с нейронными сетями. Тогда они ещё только начали появляться.

Я хотел понять, как ИИ может взаимодействовать с мозгом.

За последний год я провел ряд экспериментов на мышах и крысах. Я вживлял им имплант с ИИ.

Затем были собаки.

А сейчас наступил такой неожиданный для меня момент, когда чип я буду вживлять самому себе.

Так как последствия самой операции никто не может предугадать – такие операции ещё, кажется, не производились вовсе.

И уж тем более не понятно, во что я превращусь после того, как чип начнет встраиваться в работу моего мозга.

Поэтому я решил начать вести дневник записей.

Также я решил, что нужно рассказать краткую предысторию. Потому что без предыстории многое будет в моем дневнике непонятным.

12 марта 2023 года

Итак, в 2013 году я начал свои первые эксперименты с нейроинтерфейсами, вживляя микрочипы в моторную кору мышей. Эти крошечные устройства, размером с рисовое зерно, содержали примитивный ИИ, способный стимулировать нейроны мозга в ответ на внешние сигналы. Моя цель была амбициозной: научить мышей выполнять сложные задачи, управляя их движениями через чип. Я создал лабиринт с множеством развилок, тупиков и вознаграждений в виде кусочков сыра, чтобы проверить их поведение.

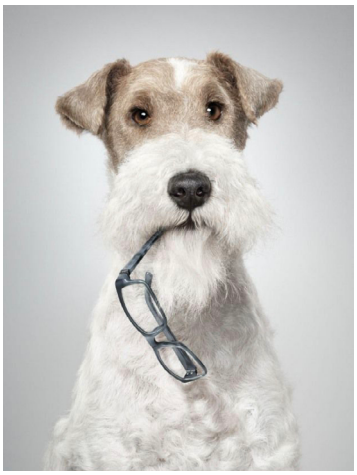
Без чипов мыши блуждали хаотично, полагаясь на инстинкты и случайные решения.

Но после имплантации всё изменилось. Чипы посылали точные электрические импульсы, направляя нейронную активность в моторной коре, и мыши двигались с пугающей точностью. Они преодолевали лабиринт за считанные секунды. Их движения были механическими: никаких лишних шагов, никаких колебаний, только прямолинейная эффективность. Я наблюдал, как они скользили по коридорам, будто крошечные автоматы, лишённые страха или любопытства.

Данные с чипов показывали синхронизированную активность нейронов, словно мозг мышей стал частью машины. Однако я заметил, что они перестали обнюхивать углы или замирать, как делали обычные мыши. Это было не просто обучение — это было подчинение. Иногда я задавался вопросом, осознают ли они свои действия, или их разум стал лишь проводником для моего кода. Но тогда я отгонял эти мысли: в конце концов, это были всего лишь мыши, а я искал границу разума.

Моя цель и сегодня остается такой – найти границу между разумом и машиной.

9 июля 2024 года



Перешёл к собакам. Их мозг ближе к человеческому.

К лету 2024 года я разработал чип второго поколения, значительно превосходящий предыдущие модели. Этот нейроинтерфейс, размером с ноготь, вживлялся в область мозга, отвечающую за принятие решений и сложное поведение.

В отличие от первого чипа, который лишь посылал импульсы для стимуляции нейронов, новый анализировал поведение мозга, корректировал команды в реальном времени и усиливал его способности.

Для своего первого эксперимента я выбрал фокстерьера со смешной кличкой Фрося — здоровую, умную собаку с развитыми инстинктами. Фокстерьеры известны своим умом и способностью быстро обучаться. Их развитые когнитивные способности делают их идеальными для экспериментов с нейронным чипом, который анализирует и усиливает мозговую активность.

После операции я поместил её в комнату с головоломками, которые обычно использовались для тестирования приматов: кубики с цветными узорами, запоры и замки, требующие последовательности действий.

К моему изумлению, за неделю Фрося освоила задачи, которые шимпанзе решали за месяц. Она манипулировала замками и запорами с точностью, будто заранее знала правильную комбинацию, а кубики складывала в нужном порядке без единой ошибки.

Данные с чипа показывали, что ИИ не просто управлял его действиями, а усиливал нейронные связи, ускоряя обработку информации.

Но знаете…

Больше всего в ходе эксперимента меня поразило то, как поменялся взгляд фокстерьера.

Теперь, когда она смотрела на меня, её глаза казались не собачьими, а почти человеческими, полными странной, пугающей осмысленности.

И ещё одно.

Я заметил, что Фрося больше не виляла обрубком своего хвоста от радости. И совсем перестала реагировать на похвалу. В ответ на все мои радостные восклицания, посвященные её успехам, она только оглядывалась на меня. Будто… Будто она хочет что-то у меня спросить.

Только не может, речевой аппарат у собак не развит.

Я начал подозревать, что чип не просто улучшал её разум.

Чип менял саму природу её сознания.

10 июля 2024 года

Сегодня встретился с доктором Еленой Соколовой, этиком НИИ. Она пришла в мою лабораторию, где я занимался с Фросей.

– Не помешаю? Могу поприсутствовать? – спросила Соколова.

– Насколько мне известно, специалисты по биоэтике имеют доступ ко всем экспериментам с животными, – ответил я.

Доволен я не был. Эти биоэтики только болтаются под ногами и, как мне кажется, никакой пользы от них нет. А вот вреда они наделать могут. Там не сломай, здесь не перегни, психика животных, смерть подопытных допустима лишь в исключительных случаях и так далее.

Елена присела в кресло у вольера. Некоторое время она наблюдала за нашими упражнениями.

– Алексей Петрович, Вы… Вы видите, как изменились глаза этой собаки? – произнесла этик после того, как Фрося подошла к заградительной сетке и стала долгим протяжным взглядом изучать биоэтика. – Я заметила это вчера, когда просматривала записи. Потому пришла сегодня посмотреть на это наяву.

– Заметил, – ответил я коротко. Я не видел причин врать. Но и развивать тему не хотелось.

– Вы не думаете, что у неё появился какой-то иной уровень сознания? – спросила она.

– Иной уровень сознания… – пробормотал я. Кажется, начинается.

– Никогда не думал об этом. Я изучаю границы между разумом и машиной, – я очень хотел побыстрее закончить этот разговор, – а разница в уровнях сознания… Я очень смутно понимаю разницу между уровнями сознания. Оно же, сознание, всегда было у этой собаки, правильно?

Елена молчала. Кажется, она поняла моё нежелание говорить на эту тему. И она решала, что делать дальше.

Биоэтик ещё некоторое время наблюдала за нами.

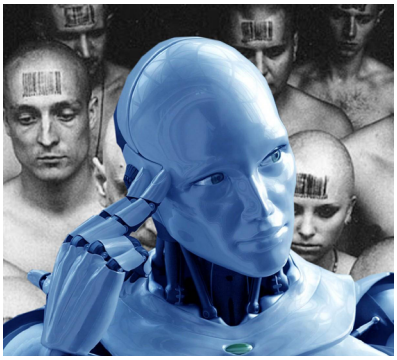
– Мне кажется, у неё совсем исчезла радость от жизни, – наконец сказала она. – Она… кажется, она стала понимать очень многое про свою жизнь.

– Что это она такое стала понимать? – я уже демонстративно не смотрел биоэтика. Я сидел за столом и копался в своих записях, хотя особенной нужды в этом не было. Я просто изображал занятость.

– Мне кажется, она поняла, для чего была предназначена природой, – ответила Соколова. – Кем стала. И кем умрет.

Я пожал плечами. Отвечать мне было нечего. Честно сказать, я и сам заподозрил что-то такое. Но развивать эту тему мне не хотелось вовсе.

У моего эксперимента были совсем другие цели.



15 сентября 2024 года

Фрося умерла.

Перестала есть, пить, легла в угол вольера и затихла. Умерла подозрительно быстро, не прошло и двух суток. Я человек не сентиментальный, но её мне был очень жаль. И даже не потому, что у меня прервался эксперимент. Нелепая смерть хорошего человека, так сказать. И ещё её эти глаза. Которые так изменились во время эксперимента.

Так они и стоят передо мной до сих пор.

После Фроси у меня было ещё шесть собак разных пород.

Результаты их обучаемости были ещё более потрясающими, чем у Фроси. Вот только… Вот только и погибали они много быстрее первого фокстерьера. И чем дольше обучался чип, тем быстрее дохли мои подопытные.

После смерти шестой собаки я понял, что «в консерватории» нужно срочно что-то подправлять.

Тем более, что биоэтик Соколова, конечно же, тоже не сидела без дела.

Благодаря её докладным меня уже замучили проверками. И если бы не заступничество замдиректора НИИ по инновациям Зубова, мои исследования уже давно бы прикрыли.

По поводу последней собаки мне позвонил сам Зубов с коротким «разберитесь, Ковалев».

После этого я приступил к тщательному анализу данных, поступающих с нейрочипов, имплантированных в префронтальную кору головного мозга собак, чтобы выявить возможные причины их внезапного отказа от пищи и заметного снижения поведенческой активности, включая апатию и отсутствие интереса к привычным стимулам.

Нейронная активность, записанная в последние дни их жизни, показывала аномалии в мозге, схожие с теми, что наблюдаются у людей с клинической депрессией. К примеру, в префронтальной коре мозга резко снижалась активность дофаминовых путей, ответственных за мотивацию и удовольствие.

Эти изменения указывали на глубокое эмоциональное угнетение, словно собаки утратили интерес к жизни.

Я предположил, что ИИ, усилив их когнитивные способности, дал им способность осознавать своё положение. А положение было незавидным, если его вдруг осознать — замкнутое пространство клеток, бесконечные тесты, отсутствие свободы. Разум собак, обострённый чипом, начал анализировать условия существования.

И они, кажется, пришли к выводу, что такая жизнь лишена смысла.

16 сентября 2024 года

Соколова снова появилась в лаборатории после того, как я ей позвонил.

– Я исправлю код, Елена, – сказал я биоэтику. – Я нашел основной баг. Вы были правы. Они, кажется, стали осознавать, что … что находятся не в том месте и занимаются не тем, для чего рождены природой.

– И что же? Вы хотите дать им смысл жизни? – Спросила Соколова. – Вы что же… Думаете, что вы – Бог?

– Я думаю, что я доцент кафедры нейробиологии Новосибирского НИИ когнитивных технологий, – улыбнулся я. – Дайте мне спокойно поработать несколько месяцев и вы увидите результат.

– Вы слепы, как и многие ученые, Алексей Петрович, – сказала биоэтик. – Вы прете против природы танком, только не осознаете, что танком природу ни расстрелять, ни задавить.

– Да поймите же, – сказал я, хотя прекрасно знал, что ничего такого она понять не сможет. – Это всего лишь гиперактивность префронтальной коры. Я добавлю ограничения в код, чтобы подавить ненужные реакции.

– И получите новых мертвых собак, – сказала на это Соколова. – Просто умрут они не от депрессии, а от чего-то другого.

3 ноября 2024 года

Новый алгоритм: ИИ не инициирует самоуничтожение, он стимулирует выживание.

Новая партия собак. Но теперь они обучаются не по очереди, а все вместе.

Шесть лабрадоров решают задачи, синхронизируют движения, будто стая с общей целью.

Новый подэксперимент длится уже третью неделю, и нет ни намека на суицидальные настроения.

Моя радость была беспредельной, пока не случилось что-то мало вообразимое.

Через три недели эксперимента все шесть лабрадоров сбежали.

Расследование, которое потом проводилось, показало слаженную и профессиональную работу команды. Так работают команды, ну скажем, грабителей. Грабителей-людей.

Лабрадоры Алеф, Бет и Гимел за одну ночь по очереди расшатали решетку вентиляции в общем вольере. Зубами и лапами они работали над ней, постепенно ослабляя винты, на которых держалась решетка. Они не грызли её – видимо, они понимали, что это было бесполезно. Они её именно разобрали.

Собаки работали в полной тишине, избегая датчиков звука, которые были установлены в вольере для контроля.

Затем все шесть собак забралась в вентиляцию, пробрались в технический коридор, где Алеф намеренно вызвал короткое замыкание, отключив резервную систему сигнализации. Алеф так и остался лежать прямо на проводах, которые замкнул собой.

После этого оставшихся пять лабрадоров — Бет, Гимел, Далет, Хе и Вав — продолжили свой путь через тёмный технический коридор.

Их чипы, анализируя данные с нейронных сетей, позволяли им ориентироваться в здании лучше, чем в нём ориентировались те, кто его построил.

Бет, теперь взяв на себя роль лидера, остановилась у распределительного щита, где собаки заметили мигающие индикаторы системы безопасности. Используя зубы, она осторожно выдернула ключевой провод, отключив датчики движения в коридоре, не вызвав при этом общей тревоги. После этого собаки вышли в общий коридор и теперь проявить себя настал черед Бет.

Он разогнался и прыгнул в оконный проем. От веса и скорости собаки оба стекла окна посыпались. Рама была огромной, два на три метра. Потому даже используя свой чип, Бет не смог избежать острых порезов. Следователи нашли Бет примерно в трёхстах метрах он места побега. По кровавому следу.

Дальше была Гимел.

Стая как будто знала все свои имена в алфавитном порядке и договорилась в случае опасности жертвовать собой именно так, по алфавиту.

Сначала они соорудили около забора нашего НИИ конструкцию, которая позволяла собакам перемахнуть забор. Гимел прыгнула с неё и повисла на колючей проволоке, чтобы остальные смогли выбраться, прыгая сначала на её тело, а потом уже вниз.

С «колючки» Гимел сползти ещё смогла, а вот идти дальше – уже нет.

Когда я со следователями проходил каждый труп собаки, в глазах у меня стояли слёзы. При виде Гимела я извинился перед группой и отошел в сторону. Мужчины не плачут. В присутствии других мужчин.

Далет, Хе и Вав ушли в ближайший лес. Их ищут до сих пор.

4 ноября 2024 года

– Для людей собаки гарантированно безопасны, – говорил Зубов кому-то по телефону, когда я зашел к нему в кабинет. – Вы же понимаете, что если у них хватило ума сбежать из нашего НИИ, то хватит ума понять..., да... так вот, они с гарантией сто процентов сообразят, что усилия по их поимке увеличатся в десятки раз, в случае нападения или угрозы для людей. Ищем, Иван Палыч, конечно, ищем. Понял. И вам.

Зубов положил трубку. Немного помолчал. Я присел напротив. Мне показалось, что это будет не короткий разговор.

– Мы замораживаем твой эксперимент, Леша, – сказал Зубов. – На неопределенное время. Слишком уж из всего этого получается что-то непредсказуемое. Готовь отчет о проделанной работе.

– Есть, – ответил я. Потом внезапно для себя понял, что разговор уже состоялся, встал и вышел из кабинета.

5 ноября 2024 года

Что ж. Предыстория окончена. Начинается история.

Я не успел создать семьи, у меня почти не было друзей, хобби и развлечений.

У меня был только мой эксперимент.

И теперь его хотят заморозить. На неопределенное время.

Если человек знает «зачем», он выдержит любое «как».

Конечно же, я не планировал этого делать. Но теперь меня приперли к стенке. И я вынужден. Вынужден пойти на преступление. Я прекрасно знаю, что вживление чипа любому homo – это уголовно наказуемое преступление. Я подписывал соответствующие бумаги при устройстве в наше НИИ.

Для операций по вживлению чипов у меня есть робот-хирург.

Я запустил его, настроил, сам лег на операционный стол.

Мне казалось, что холод металла порой проникал сквозь тонкую ткань накидки, пока робот-хирург с жужжанием калибровал свои инструменты.

Чип второго поколения, тот самый, что сделал собак разумными и обрёк их на страдания, находился в манипуляторе — я модифицировал его, чтобы усилить функции, которые отвечают за познание, но подавить эмоциональные всплески.

Операция прошла безупречно: робот вживил чип в мою префронтальную кору, и я почувствовал лёгкий гул, будто ток пробежал по нейронным связям моего мозга.

14 марта 2025 года

Мои мысли стали острыми, как лезвие — я анализирую массивы данных за секунды, вижу закономерности в хаосе, решаю задачи, которые раньше требовали месяцев.

Но эмоции начали угасать.

Я смотрел на фото матери, и её улыбка казалась лишь набором пикселей.

Я получил письмо от старого коллеги, который поздравлял меня с днём рождения, но текст сейчас казался лишь последовательностью символов, лишённой значения, а само понятие этого праздника – странным, неэффективным ритуалом.

Как-то, проходя мимо парка, я услышал детский смех. Но вместо того отклика, которое когда-то вызывал звук детского смеха – тоже, в общем, не слишком, но хоть сколько-нибудь эмоционального – мой мозг сейчас просто разложил звуки на частоты и амплитуды, как данные с осциллографа.

Кажется, я сам превращался в осциллограф.

15 марта 2025 года

Вчера был мой день рождения. Ели торт. Вкус его мне казался просто химической комбинацией сахара и жиров, не вызывающей ничего, кроме анализа калорий.

Елена Соколова, которую я по настоянию Зубова тоже пригласил вчера на свой день рождения, подошла ко мне после того, как торт был съеден и все разошлись.

– Алексей Петрович, с вами что-то не так, кажется? – спросила она.

– Что же со мной не так? – спросил я.

– Вы стали какой-то совсем нелюдимый... Вас совсем не интересует ничего из жизни НИИ, вы даже тему для новых исследований не предоставляете уже который месяц.

Я действительно не видел больше никакого смысла в том, чтобы что-то там разрабатывать для НИИ и тянул, как мог, с разработкой новой темы.

На работу я ходил по какой-то, кажется, привычке.

На рабочем же месте я занимался в основном тем, что сидел и размышлял.

С моими новыми способностями размышлять было радостно и легко. До тех пор, пока я не упирался в тупик собственной смерти или конечности всего сущего вообще.

– А раньше у меня было много друзей в НИИ, Елена? – спросил я Соколову. – И я принимал участие в рисовании стенгазеты?

– Нет, не принимали...– кажется она думала о чем-то своем.

Мы помолчали.

– Вы знаете, что Далет и Хе были вчера застрелены в Заельцовском бору? – внезапно спросила Соколова.

Я внимательно на неё посмотрел. Она не менее внимательно смотрела на меня.

Тогда, когда я стоял над трупом лабрадора Гимель, израненной колючей проволокой, я заметил очень похожий на теперешний внимательный взгляд биоэтика Соколовой.

И тут я все понял. Она ждала от меня проявления эмоций. Кажется, они что-то подозревают.

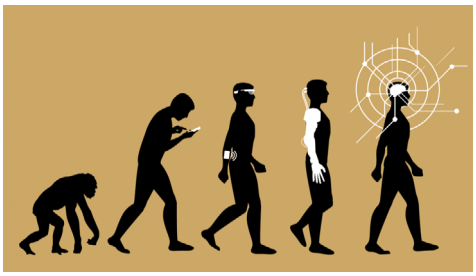
– Врете, – сказал я. – Во-первых никто не стал бы их стрелять – собак можно просто усыпить. Во-вторых, я уверен на 98,6 процента, что собаки Даллет, Хе и Вав держатся вместе. Все втроем, стаей. И почему тогда застрелили только двоих?

Тут я осекся.

Да не отсутствия эмоций она от меня ждала, черт побери!

А, скорее всего, вот этих вот 98,6 процента.

И теперь стало совсем уж очевидным, зачем Зубов настоял на том, чтобы Соколова пришла на мой день рождения.



Я думаю, он давно заподозрил, что я внедрил себе чип.

Но память робота-хирурга я очистил, чип списал, все следы самого внедрения чипа на моей голове уже зажили.

И как-то доказать то, что я совершил преступление по внедрению себе чипа, по косвенным данным уже было нельзя. Меня, конечно, можно было обследовать по-другому, но для этого нужны были хоть какие-то основания.

Поэтому Зубов вызывал меня на разговоры уже раза четыре за последних пол месяца. Предлоги были разными – в основном говорили о разработке новой темы. Но я прекрасно понимал, что с Зубовым нужно держаться ну очень осторожно. И, видимо, ничем себя не выдал. Потому Зубов прислал ко мне Соколову.

И того, чего не получилось у Зубова. Сейчас, кажется, получилось у Елены Николаевны.

– Ну не знаю, – ответила она. – За что купила, за то и продаю.

Она говорила совсем уж какие-то глупости и стало совсем очевидно, что чушь про Даллет и Хе была лишь тем, что ей сказал передать мне Зубов. Но, кажется, биоэтик так и не поняла, что точность прогноза в 98,6 процентов в данной ситуации может дать только машина.

Зато это прекрасно поймет Зубов.

И тогда меня ждет уже совсем другая проверка.

Соколова ушла, и её шаги затихли в пустом коридоре. Я остался один в комнате, где ещё витал приторный запах ванили от недоеденного торта. Догоревшие свечи оставили на столе лишь лужицы воска, и я смотрел на них, впервые за месяцы позволив себе остановить бесконечный поток расчётов, который чип гнал через мой мозг.

16 марта 2025 года

Внезапно в голове сами собой всплыли три цифры. Три очень важных цифры.

98,6%, 63,2%, 26,4%.

Про первое число вы все знаете.

Второе – вероятность моего разоблачения.

Третье – вероятность того, что я могу сейчас выйти за двери нашего НИИ и раствориться, исчезнуть среди миллионов россиян. И что меня никто никогда не найдет.

А что будут искать – сомнений не было.

Вероятности, проценты, холодные данные – они заполняли всё, вытесняя то, что когда-то было мной.

Я тяжело встал, опершись на стол обеими ладонями и понял, что влез ладонью в липкое пятно от крема.

И тут я вспомнил, как в детстве, ещё мальчишкой, украдкой слизывал такой же крем с ложки, пока никто не видел.

Тогда я не знал ничего о химическом составе сахара или калорийности жиров. Тогда я просто радовался этому мгновению – простому, глупому, живому.

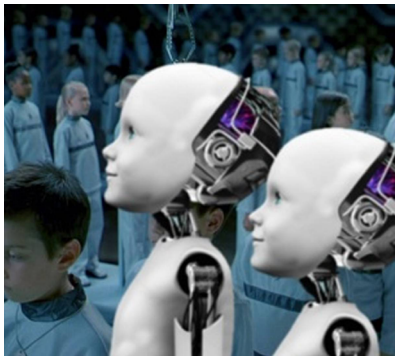
Чип отнял у меня это.

Он сделал меня быстрее, умнее. Но украл вкус.

Вкус к жизни.

Разговоры с коллегами стали для меня бессмысленным шумом, обсуждения проектов — пустой последовательностью фраз, музыка стала лишь последовательностью нот и ритмов, не трогающих душу, закат — всего лишь спектром света, разложенный на длины волн.

Я понял, что этот торт, этот липкий след крема на ладони — он стал триггером.



Я вдруг понял, что больше не хочу быть машиной, которая лишь анализирует.

Я вдруг понял, что уже даже забыл главную идею своего эксперимента – найти границу между разумом и машиной.

Я эту границу просто незаметно пересек.

И сейчас – может быть ненадолго – вернулся обратно.

И ещё я понял, что уже не проводил эксперимент. Я просто захотел стать сверхчеловеком.

Но за все в этом мире надо платить. И платить за сверхчеловечность пришлось человечностью обычной.

Зубов со своими подозрениями и ловушками не был моим главным врагом. Врагом себе был я сам – тот, кто решил, что может перехитрить природу.

Я дотронулся до шрама за ухом, где сидел, как паразит, чип. И переписывал мои мысли.

Решение родилось не из расчётов, а из чего-то глубинного, что чип ещё не успел стереть. Я достану из своей головы чип.

Я больше не хочу быть сверхчеловеком. Оставьте мне моё. Человеческое.

Операция будет сложной – новый робот-хирург, тайные приготовления, ещё более тщательная маскировка следов. Но я сделаю это.

Не ради спасения от Зубова или страха разоблачения. А ради себя. Ради того, чтобы снова почувствовать вкус крема, услышать смех и не разложить его на амплитуды, ради того, чтобы работа снова обрела смысл.

Я взял со стола минералку, плеснул немного на кисть, стал стирать салфеткой крем и тут меня озарило.

Мой мозг. Это просто мой мозг, учуявший опасность разоблачения. И всё понявший про вот эти вот вероятности – 63,2 процента и 26,4 процента – мой мозг победил, одолел чип робота. И наружу, минуя всякие нейронные сигналы, что подавал чип, поперла снова суть человека, который просто хочет жить. Жить, а не сидеть в тюрьме!

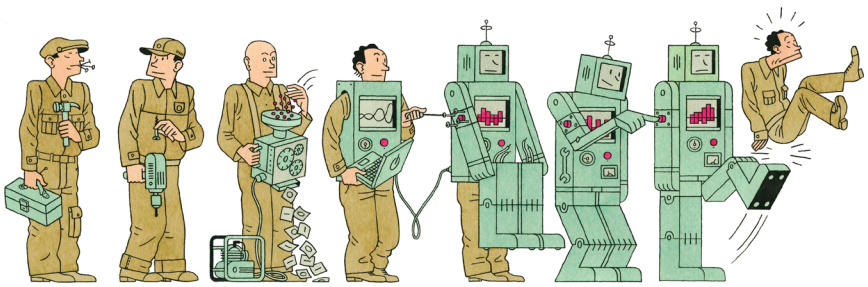
Я дотер крем с руки и расхохотался. Потом взял со стола шампанское, открытое, но почти не тронутое, налил себе полный фужер, поднял его и понял, что у меня теперь есть тост.

– Шалишь, брат, – с задором сказал я неведомо кому в темноту комнаты. –Шалишь! Человек все ещё – а может быть и всегда! – будет сильнее. Жизнь, брат! Живое всегда будет побеждать мертвое, дружище! Аминь!

Я выпил фужер, ляснул его о пол, фужер брызнул тысячей брызг.

Потом я взял со стола уже всю бутылку шампанского, развернулся и пошёл к окну. Ночной город мерцал, как и прежде, но теперь я видел в нём не поток данных с чипа, а обещание.

Обещание вернуть себя.



Рисунки заимствованы из общедоступного источника Интернет, не содержащего ссылок на авторство.

Приглашаем авторов к участию в журнале «Вестник современных цифровых технологий»

ИНФОРМАЦИЯ ДЛЯ АВТОРОВ

Редакция принимает материалы статей, соответствующие тематике журнала, содержащие новые научные результаты, не опубликованные ранее и не предназначенные к публикации в других печатных или электронных изданиях. Проводится независимое внутреннее рецензирование. Статьи в журнале публикуются бесплатно (объем – до 15 тыс. знаков), после получения одобрения Редакционного совета.

Для опубликования статьи в редакцию журнала необходимо направить по адресу *a.shcherbakov@c3da.org, a.ryazanova@c3da.org* **следующие материалы в электронном виде:**

• рукопись статьи в DOC- и PDF-форматах;
• иллюстрации, предоставленные также и отдельными файлами в формате JPG или PNG с разрешением 300 dpi;
• сведения об авторах, содержащие фамилию, имя, отчество, ученые степень и звание, должность, место работы, контактные телефоны или E-mail;
• англоязычную информацию, содержащую название статьи, ФИО авторов, аннотацию и ключевые слова;
• редакция может запросить экспертное заключение о возможности публикации статьи в открытой печати.

ПОСЛЕДОВАТЕЛЬНОСТЬ МАТЕРИАЛОВ ДЛЯ ПУБЛИКАЦИИ:

1. шифр УДК (см. Справочник УДК) в левом верхнем углу;
2. название статьи (полужирным шрифтом по центру) не более 12 слов;
3. инициалы и фамилия автора (полужирным шрифтом по центру), к каждому автору- сноска, содержащая ученое звание, должность, название организации (без сокращений), e-mail;
4. Аннотация, излагающая суть работы и полученные результаты (5-7 строк);
5. ключевые слова (8-10 слов);
6. англоязычная информация по статье (по пп.2-5)
7. текст статьи с учетом указанных далее требований к его оформлению;
8. список литературы, оформленный по ГОСТ Р 7.0.5-2008.

Статья должна быть структурирована, т.е. должна включать разделы с названиями, кратко и точно отражающими их содержание, в том числе:

• введение, содержащее обоснование актуальности и краткий обзор проблематики;
• четкую постановку задачи исследования;
• описание метода решения задачи исследования;
• прикладную интерпретацию и иллюстрацию полученных результатов исследования;
• заключение, включающее обобщение и указание области применения полученных результатов, не повторяющее аннотацию и не ограничивающееся простым перечислением того, что сделано в работе.

С детальными требованиями к рисункам, таблицам, формулам, списку литературы, а также с примерами оформления статьи можно ознакомиться на странице Вестника http://c3da.org/journal.html.

Приглашается к сотрудничеству редактор для работы в редакции журнала по совместительству. Просьба направлять резюме по электронному адресу *accda@c3da.org, info@c3da.org*

ТРЕБОВАНИЯ К РЕДАКТОРУ:

• отличное знание русского языка;
• свободное владение ПК, в том числе специальными текстовыми и графическими программами;
• опыт работы в издательстве.

Высшее техническое образование и знание английского языка являются существенными преимуществами.

ОБЯЗАННОСТИ

Редактор:

• редактирует рукописи, принятые к изданию;
• оказывает авторам необходимую помощь по улучшению структуры рукописей, выбору терминов, оформлению иллюстраций;
• проверяет, насколько учтены авторами замечания по доработке, предъявленные к рукописям;
• подписывает отредактированные рукописи в печать.