

СОЗДАНИЕ КИБЕРПОЛИГОНА: БЛОК НАВИГАЦИИ ПО ТЕХНОЛОГИЯМ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА И МАШИННОГО ОБУЧЕНИЯ

А.Л. Сердечный, А.Т. Труфанов, А.А. Карданов

В статье представлены результаты разработки блока навигации по технологиям искусственного интеллекта и машинного обучения, выполненной в рамках создания киберполигона. Блок базируется на программной реализации интерактивной информационной карты «Технологии искусственного интеллекта и машинного обучения». Информационная карта систематизирует сведения о задачах, методах и моделях машинного обучения, а также наборах данных и научных публикациях, которые размещаются на информационном ресурсе Papers With Code.

Информационная карта раскрывает сложную структуру взаимосвязей между более чем 3 тыс. задач машинного обучения, 34 тыс. научных публикаций, посвящённых их решению, а также 7 тыс. наборов данных, 2 тыс. методов и алгоритмов машинного обучения, 8 тыс. моделей, упоминаемых в таких статьях. В ходе исследования информационной карты могут быть получены знания о современном состоянии развития технологий искусственного интеллекта, что способствует более глубокому пониманию основных тенденций в данной области, а также повышает эффективность поиска по рассматриваемой теме.

Также важным практическим значением разработанной информационной карты является возможность использования её для совместной работы по темам в области искусственного интеллекта. В качестве примера построен слой, на котором показаны области размещения и взаимосвязи между задачами, моделями и наборами данных, используемых при обеспечении информационной безопасности и защите информации.

Блок навигации построен с использованием web-технологий и доступен в тестовом режиме на портале CyberMaps.ru.

Ключевые слова: киберполигон, информационная карта, информационное картографирование, искусственный интеллект, машинное обучение.

Введение

Одним из ключевых направлений использования киберполигона, создаваемого на кафедре систем информационной безопасности Воронежского государственного технического университета, является исследование технологий искусственного интеллекта в контексте обеспечения информационной безопасности и защиты информации. Их стремительное развитие требует от исследователей постоянной актуализации своих знаний, что с каждым годом становится все труднее обеспечить из-за лавинообразного появления новых публикаций, задач, методов и терминов в области машинного обучения.

Частично упростить анализ всё возрастающего объёма сведений помогают специализированные автоматизированные поисковые системы [1,2,3] и

информационные ресурсы, агрегирующие публикации по теме [4,5]. Также сами технологии искусственного интеллекта способствуют появлению новых средств, повышающих эффективность исследователей. Например, ChatGPT [6] позволяет в режиме вопрос-ответ быстро получить доступ к нужной информации, однако опыт работы с подобными инструментами показывает наличие дополнительных проблем, связанных с проверкой полноты и достоверности предоставляемых данных. Кроме того, поисковые и экспертные системы ориентированы, в основном, на отдельных лиц и не так эффективны в коллективной работе, как, например, wiki-порталы [7] или платформы для совместного картографирования [8], которые предоставляют единое информационное пространство и удобные инструменты множеству авторов для коллективной работы над темой.

В условиях колоссального роста информации особенно остро стоит проблема обеспечения широты и полноты представление накапливаемых знаний о технологиях искусственного интеллекта. Для её решения в рамках создания киберполигона была поставлена задача разработки навигационного блока по технологиям искусственного интеллекта и машинного обучения. Результаты её решения представлены в настоящей статье.

Разработка навигационного блока осуществлялась на базе исследования, проводимого с целью систематизации сведений о технологиях искусственного интеллекта и машинного обучения для последующего определения исходных данных, которые могут быть использованы для решения задач в области обеспечения информационной безопасности и защиты информации.

Исследования проводились на основе метода информационного картографирования [9], так как его применение для решения задач, связанных с экспертным анализом, систематизацией и представлением большого объёма данных показало свою эффективность [10, 11, 12]. Суть метода заключается в построении и анализе информационной карты, с помощью которой обеспечивается представление всего многообразия сведений о предметной области, что позволяет исследователю одновременно наблюдать как объект исследования, так и контекст, в котором он находится. При этом за счёт масштабирования информационной карты обеспечивается возможность детализации структуры объекта исследований.

Результаты, получаемые в ходе исследования, размещаются на информационной карте, что обеспечивает единообразие их представлении. Это даёт возможность осуществления совместной работы нескольких человек, а интерактивный анализ информационной картой позволяет исследователям быстро получать доступ ко всей необходимой информации, что обеспечивает сохранение концентрации внимания, так как не требуется отвлекаться взаимодействие с другими средствами.

Исследования проводились с помощью информационно-картографической системы [13] картографирования рисков защищаемого киберпространства [14], разработанной на кафедре систем информационной безопасности (рис. 1).



Рис. 1. Основные компоненты системы картографирования рисков, использованной для построения и анализа информационной карты «Технологии искусственного интеллекта и машинного обучения (Papers With Code)» [14]

В соответствии с задачей исследования и с учётом технологии построения и анализа информационной карты [9] были выполнены следующие работы:

- сбор исходных данных о технологиях искусственного интеллекта и машинного обучения;
- формирование ландшафта информационной карты на основе собранных данных;
- разметка ландшафта информационной карты;
- поиск и анализ участков информационной карты, содержащих сведения, связанные с вопросами обеспечения информационной безопасности и защиты информации;
- программная реализация блока навигации по технологиям искусственного интеллекта и машинного обучения и включение его в состав киберполигона.

Исходные данные о технологиях искусственного интеллекта и машинного обучения

В качестве источника данных для построения информационной карты был выбран Papers With Code [4], так как он является одним из самых крупных информационных ресурсов, систематизирующих сведения о технологиях искусственного интеллекта и машинного обучения. Платформа Papers With Code [4] представляет собой базу знаний о методах, задачах, публикациях и программных продуктах в области искусственного интеллекта и машинного обучения. В 2018 году авторы проекта поставили перед собой цель устранить разрыв между академическими исследованиями в и их реализацией в коде.

Платформа [4] позволяет исследователям и разработчикам обмениваться своими результатами по заявленной теме. На данный момент Papers with Code содержит обширную коллекцию исследовательских работ (более 34 тысяч на начало 2023 года), связанных с моделями машинного обучения, наборами данных, а также задачами, решаемых в рамках проведения соответствующих исследований. Для каждой статьи имеется раздел, в котором приводятся показатели и их значения для соответствующих аналогов, описанных в рамках схожих исследований по теме. Приводятся ссылки на репозитории исходного кода и наборы данных, использованных в статье.

В результате анализа содержимого сайта определена модель исходных данных, представленная на рис. 2.

Для сбора указанных данных был использован модуль Web Scraper, входящий в состав системы картографирования рисков защищаемого киберпространства [14]. Пример страницы с разметкой собираемых данных показан на рис. 3.

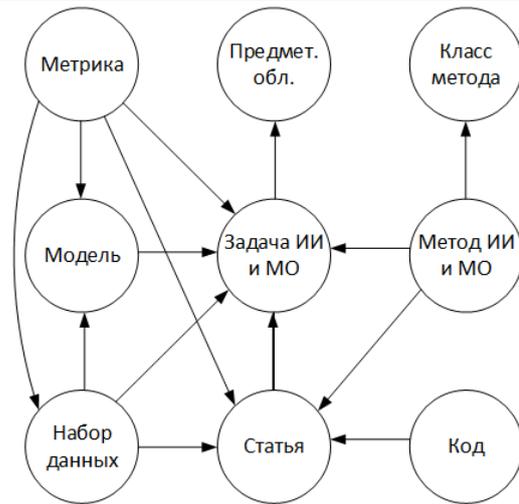


Рис. 2. Модель взаимосвязи исходных данных, собираемых с информационного ресурса Papers With Code [4]

На момент начала апреля 2023 года были собраны сведения о взаимосвязи:

- 34 987 научных статей;
- 16 предметных областей (Computer Vision, Natural Language Processing, Methodology и др.);
- 3479 задач машинного обучения (Semantic Segmentation, Machine Translation, Recommendation Systems, Time Series Analysis, Graph Embedding, Code Generation и др.);
- 8877 моделей машинного обучения (BEiT-3, BiBERT, GHRS, ResBiLSTM, DeepGG, Reflexion (GPT-4));
- 111 классов методов (Convolutions, Feedforward Networks, Attention Mechanisms и др.);
- 2034 методов машинного обучения (Transformer, Generative Models, Graph Convolutional Networks и др.);
- 7629 наборов данных (CIFAR-10, ImageNet, MNIST, KITTI, LibriSpeech и др.).

Информационная карта «Технологии искусственного интеллекта и машинного обучения (Papers With Code)»

Ландшафт информационной карты «Технологии искусственного интеллекта и машинного обучения (Papers With Code)» представляет собой граф, уложенный в двухмерном пространстве с помощью силового метода ForceAtlas2 [15].

Рельеф (структура) ландшафта существенным образом зависит от выбора

состава узлов и связей такого графа, так как различные сущности отражают действие различных факторов, определяющих расположение объектов в соответствующем пространстве карты.

Например, если рассмотреть граф связей между *научными публикациями* и *наборами данных*, то ключевую роль в расположении его узлов будет играть фактор тематики исследований, так как выбор исходных данных в первую очередь обусловлен темой. В этом случае кластеры, образованные наиболее похожими научными публикациями, образуются благодаря одинаковым наборам данных, использованных в ходе исследований. Если

же в качестве и узлов графа выбрать методы и их классы, то полученный ландшафт будет отражать лишь методологический фактор, который к тому же определяется субъективным выбором экспертов, участвовавших при классификации методов машинного обучения. Совмещение обоих вариантов позволит учесть каждый из факторов, однако, если они противоречивы, то ландшафт на базе композитного графа будет иметь большое число областей со слабо связанными объектами.

Примеры различных комбинаций, рассмотренных при выборе графа, показаны на рис. 4.

Название статьи

Аннотация

Код

Задачи

Наборы данных

Task	Dataset	Model	Metric Name	Metric Value	Global Rank	Uses Extra Training Data	Result	Benchmark
Video Captioning	MSR-VTT	GIT2	CIDEr	75.9	# 2	✓	🔄	Compare
			METEOR	33.1	# 4	✓	🔄	Compare

Модель

Метрики

Рис. 3. Пример страницы, содержащей сведения о связях научной публикации с задачами машинного обучения, наборами данных, кодом, моделями и метриками [4]

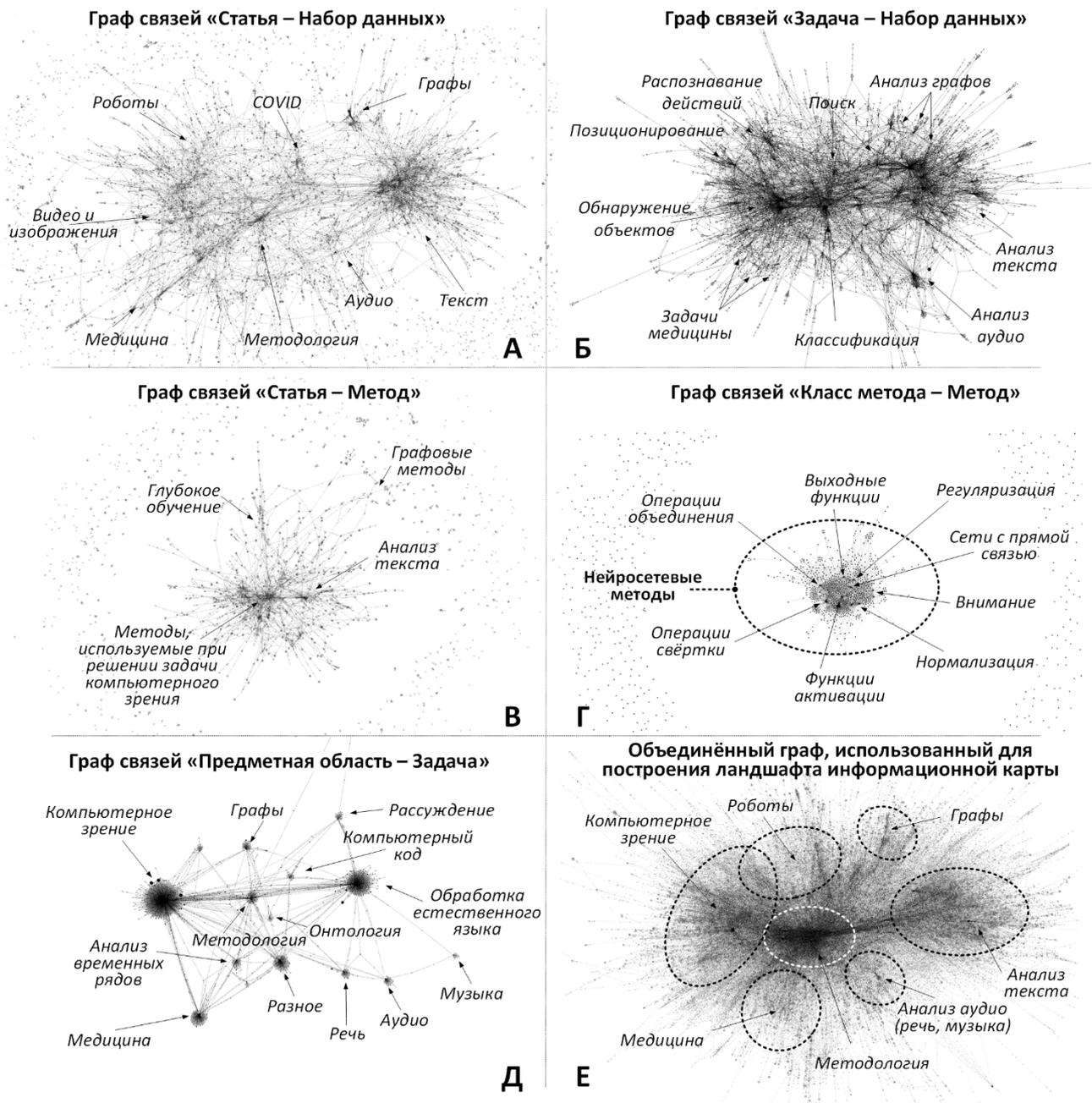


Рис. 4. Примеры возможных вариантов ландшафтов информационной карты, рассмотренных в ходе выбора состава узлов и связей графа (А – графа связей научных статей и наборов данных, Б – графа связей задач машинного обучения и наборов данных, В – графа связей научных статей и методов машинного обучения, Г – графа связей методов машинного обучения и их классов, Д – графа связей предметных областей и задач машинного обучения, Е – композитный граф, объединяющий все предыдущие варианты)

Предварительный анализ ландшафтов с целью выбора наилучшего варианта проводился с помощью программного средства Gephi, входящего в состав системы картографирования рисков защищаемого киберпространства [14].

Граф связей *научных статей и наборов данных* (рис. 4, А), имеет чётко выраженную кластерную структуру. Причём её возникновение обусловлено объективными причинами, а не субъективным выбором группы экспертов, так как кластеры образуются благодаря множеству объектов, которые напрямую не связаны друг с другом (тематика определяется не прямым указанием соответствующего тэга, а в результате косвенной связи через одинаковые наборы данных). В левой области такого графа расположены статьи и наборы данных, по теме анализа видео и изображений, в правой – анализа текстовых данных.

Из интересных особенностей можно отметить близость разряженного кластера «Аудио» к кластеру «Текст» и наличие двух обособленных кластеров публикаций в области здравоохранения (кластеры «Медицина» и «COVID»).

Кластер «Аудио» располагается ближе к кластеру «Текст», а не к «Видео», так как большую его часть составляют публикации, связанные с задачей анализа *речи*. Методы, применяемые для её решения, ближе к методам анализа текстовой информации, чем к методам машинного зрения.

Кластер «COVID» отделён от основного («Медицина») по следующим причинам. Во-первых, объявленная в 2020 году пандемия COVID-19 привела к возникновению большого потока научных публикаций и наборов данных по соответствующей теме, что для борьбы с пандемией от исследователей помимо решения задач, связанных с разработкой вакцин, требовалось еще и решение задачи автоматизированной обработки сообщений в социальных сетях (например, для выявления ареала распространения вируса). Наборы данных, используемые для решения такой задачи также используются и для анализа текстовой информации (соответствующие

публикации расположены в левой части карты).

Схожую структуру можно наблюдать и для графа «Задача – Набор данных» (рис. 4, Б), так как в нём также основную роль играет тематический фактор. При этом состав тем отличается варианта из-за большего влияния фактора экспертов. Левые и правые части ландшафта более сбалансированы, а центральная часть стала образовывать чётко выраженный кластер, связанных с решением методологических задач машинного обучения. Кластер «COVID» отсутствует, так как нет объективных задач машинного обучения, связанных со спецификой именно данного заболевания.

Методологический фактор ещё в большей степени проявляется на графах «Статья – Метод» (рис. 4, В), «Класс метода – Метод» (рис. 4, Г), «Предметная область – Задача» (рис. 4, Д). Причём на графе «Класс метода – Метод» действие тематического фактора практически не проявляется.

В результате анализа рассмотренных ландшафтов принято решение использовать композитный граф (рис. 4, Е), объединяющий все рассмотренные сущности, так как в таком графе представлены сразу все объекты, а влияние факторов сбалансировано. Скелет композитного графа образуют взаимосвязи научных публикаций и задач машинного обучения.

Разметка ландшафта информационной карты, построенного на базе уложенного в двумерном пространстве графа связей научных публикаций, задач и методов машинного обучения, наборов данных, классов методов и предметных областей проводилась в несколько этапов. Была проведена кластеризация графа при помощи Лейденского алгоритма [16].

Для каждого кластера эксперты проводили анализ наиболее влиятельных (имеющих наибольшее значение PageRank) узлов и плотность их расположения, после чего в программе QGIS, входящей в систему картографирования рисков защищаемого киберпространства [14] осуществлялось нанесение границ соответствующей области. При этом необходимо отметить, что такие

границы используются лишь для обозначения зон наибольшего влияния, где вероятность нахождения объекта соответствующего типа больше чем в других областях (в области, очерченной границами, могут присутствовать объекты других классов, но вероятность их нахождения существенно ниже по сравнению с вероятностью появления в своей области).

На завершающем этапе проводился анализ близлежащих кластеров и определялись границы объемлющих областей, образующие разметку слоя для более мелкого масштаба карты (рис. 5).

В результате генерализации данных была сформирована многослойная масштабируемая информационная карта (рис 6). На определённом уровне отображаются только те объекты, которые требуются для решения соответствующей задачи.

Так, например, на слое самого мелкого масштаба (1 к 32) отображались лишь наиболее крупные тематические области («Текст», «Аудио», «Роботы», «Графы», «Методология», «Компьютерное зрение», «Изображения», «Медицина»). Такой

уровень детализации необходим для самого общего представления предметной области и взаимосвязи основных направлений искусственного интеллекта и машинного обучения. Соответствующее представление подходит для задач общей навигации и используется для «миниатюры», с помощью которой может быть осуществлено быстрое перемещение в любую часть карты, а также с помощью выносок показано расположение детализированных областей.

Для более углубленного изучения взаимосвязей между темами используются масштабы 1 к 16 или 1 к 8. На слое масштаба 1 к 16 кластеры представлены в виде соответствующих областей, а масштаба 1 к 8 – в виде узлов графа, обозначающих задачи машинного обучения. Более мелкий масштаб позволяет исследовать объекты конкретных тематических кластеров.

За счёт генерализации данных и масштабирования преодолевается проблема отображения избыточного числа объектов, приводящего к перегрузке внимания исследователя, а возможность быстрого перемещения между уровнями и участками карты снижает его когнитивную нагрузку.

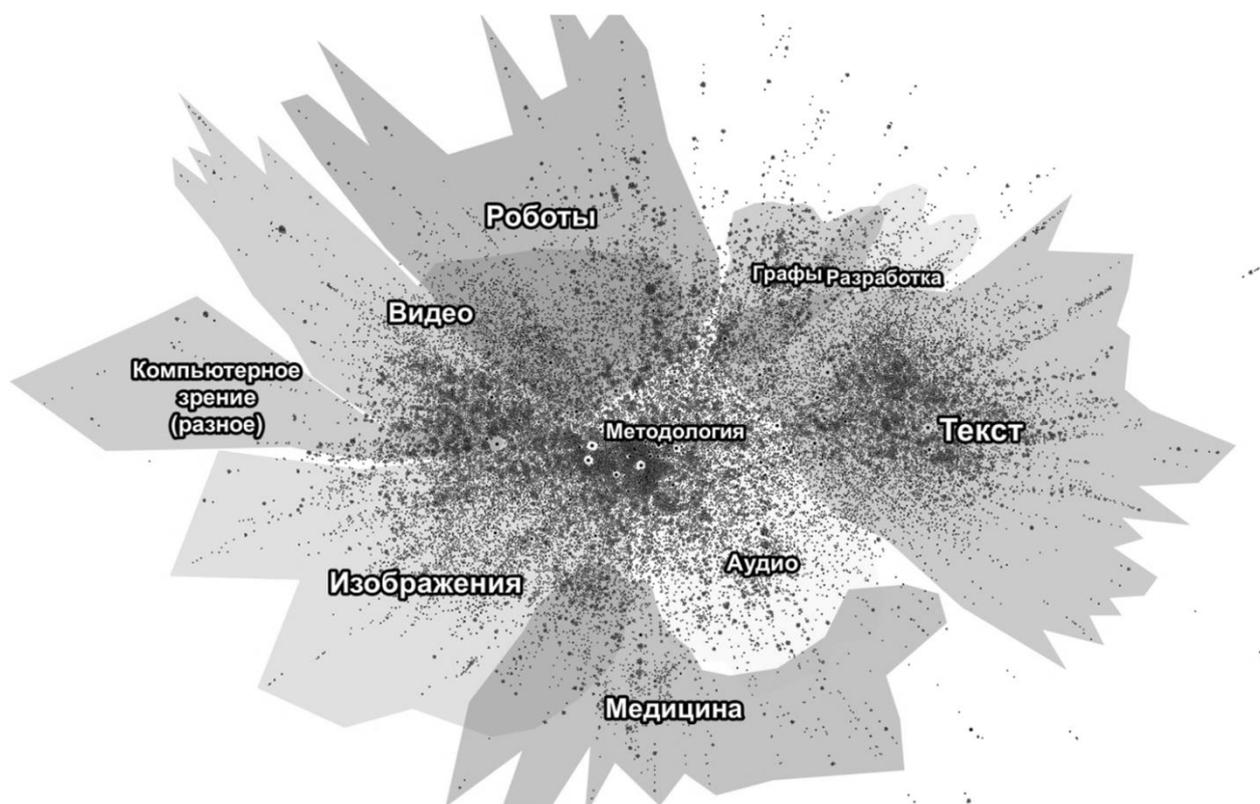


Рис. 5. Информационная карта «Технологии искусственного интеллекта и машинного обучения (Papers With Code)» с нанесённой на ней базовой разметкой

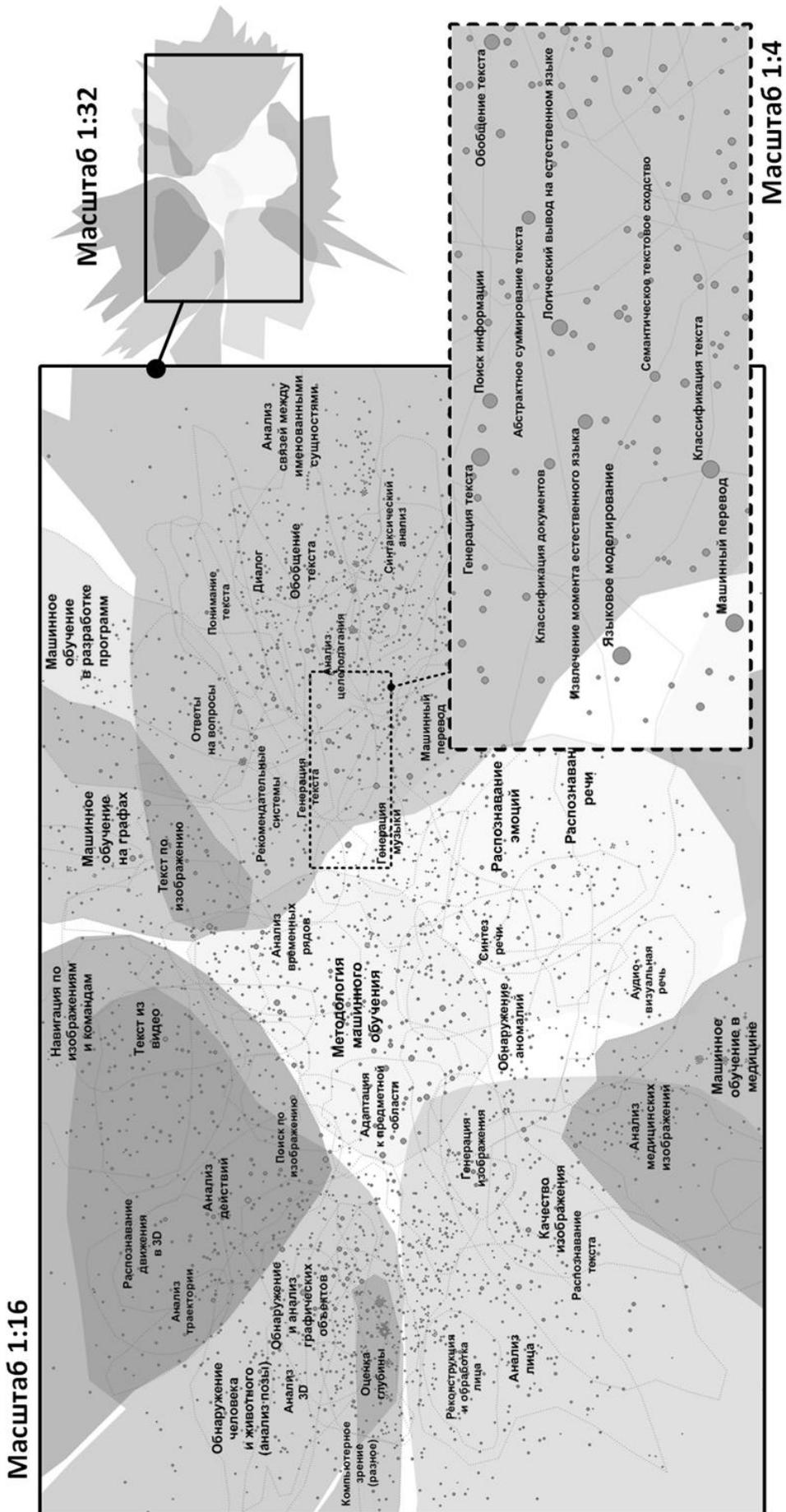


Рис. 6. Пример масштабирования информационной карты «Технологии искусственного интеллекта и машинного обучения (Papers With Code)»

Примеры практического использования информационной карты

С целью совместного использования информационной карты «Технологии искусственного интеллекта и машинного обучения (Papers With Code)» для решения практических задач в области обеспечения информационной безопасности и защиты информации в рамках работ по созданию киберполигона был разработан блок навигации по технологиям искусственного интеллекта и машинного обучения. Структура данного блока показана на рис. 7.

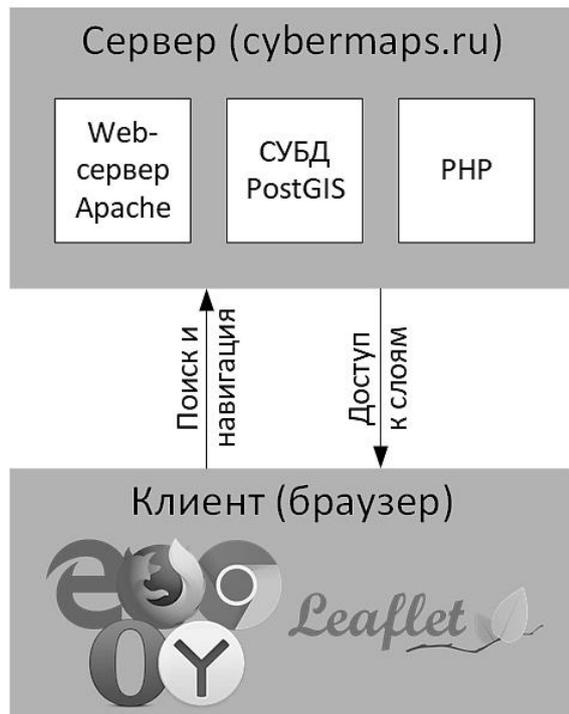


Рис. 7. Структура блока навигации по технологиям искусственного интеллекта и машинного обучения

Блок построен на базе web-технологий и реализует клиент-серверную архитектуру. Серверная часть размещена на хостинге в сети Интернет с целью предоставления повсеместного доступа к слоям информационной карты из любой точки мира. Слои информационной карты хранятся в геопространственной СУБД PostGIS, а логика обработки запросов от пользователей реализуется с помощью сценариев, написанных на языке PHP. Клиент может загрузить интерактивную карту с помощью обычного браузера. За её отображение отвечает JavaScript-библиотека Leaflet. В ходе интерактивного взаимодействия с

картой решается одна из основных задач – получение знаний о текущем состоянии в области технологий искусственного интеллекта и машинного обучения.

Другим важным направлением практического применения разработанного модуля является задача научного поиска наборов данных и результатов исследований по теме. Для этой цели был сформирован отдельный слой (рис. 8), на котором отображены лишь те объекты, которые связаны с решением следующих задач в области обеспечения информационной безопасности и защиты информации:

- защита от деструктивного контента;
- защита от вредоносных программ;
- обнаружение уязвимостей;
- расследование компьютерных преступлений;
- криптографическая защита и стеганография;
- атаки на модели машинного обучения.

Соответствующие наборы данных, научные публикации, модели и задачи машинного обучения локализованы в соответствующих областях. Так, например, объекты, связанные с задачей выявления деструктивного контента закономерно разделены на две группы (рис. 9). К первой группе относятся объекты, связанные с применением методов анализа изображений и они расположены в левой части информационной карты. В правой части карты находятся публикации, наборы данных, задачи, решаемые путём анализа текстового содержания ложных сообщений («дипфейков»).

На рис. 10 показан участок информационной карты, в котором размещены объекты, связанные с задачами противодействия вредоносным программам. Исследуя соответствующие области, можно обнаружить объекты, в названиях которых нет ключевых слов, по которым осуществлялся первоначальный поиск. Благодаря наблюдению контекста, в котором расположены такие объекты у исследователя формируется целостная картина о предметной области.

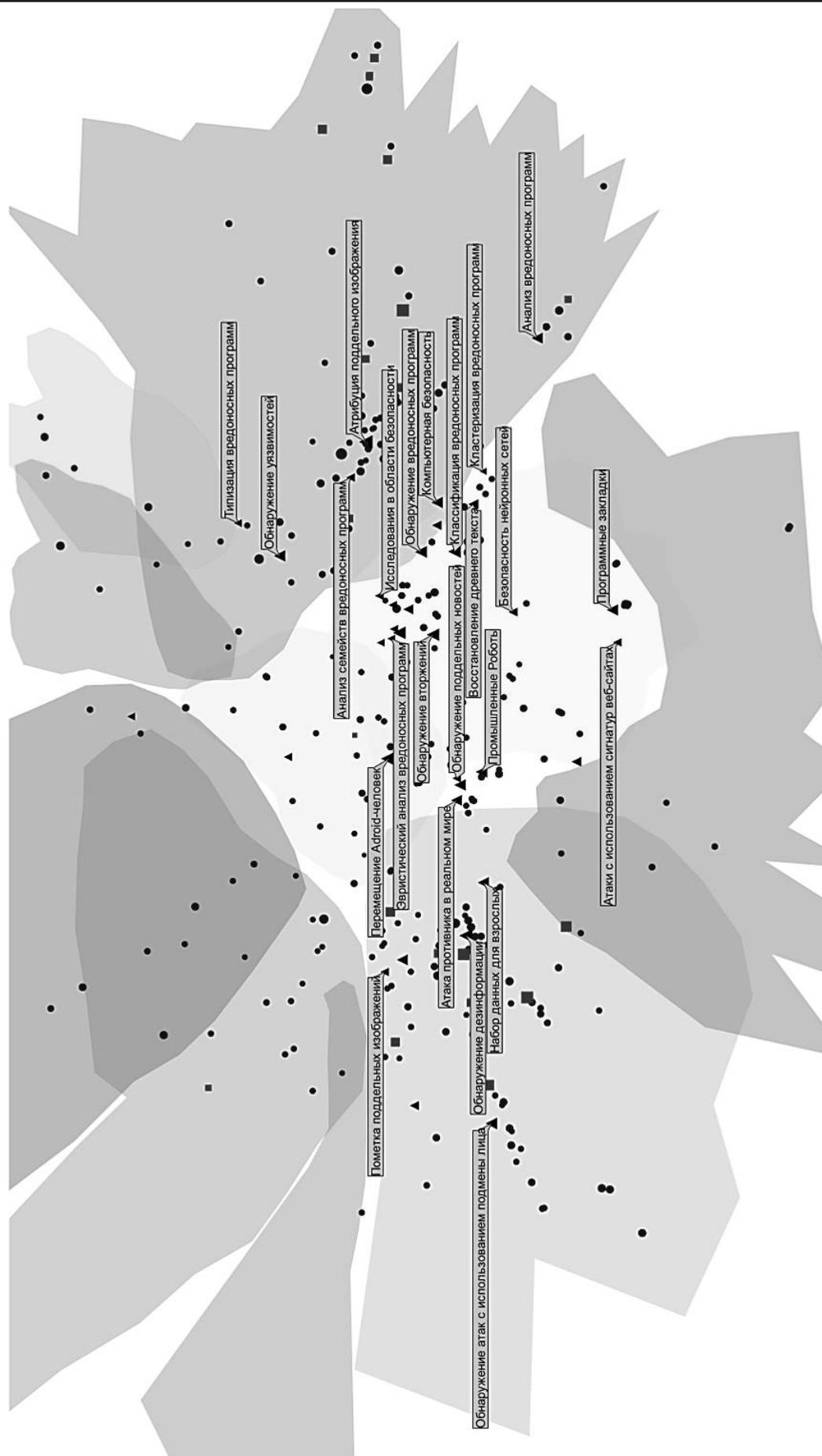


Рис. 8. Слой объектов информационной карты «Технологии искусственного интеллекта и машинного обучения (Papers With Code)», связанных с решением задач в области обеспечения информационной безопасности и защиты информации (треугольник – задача машинного обучения, квадрат – набора данных, круг – научная публикация)

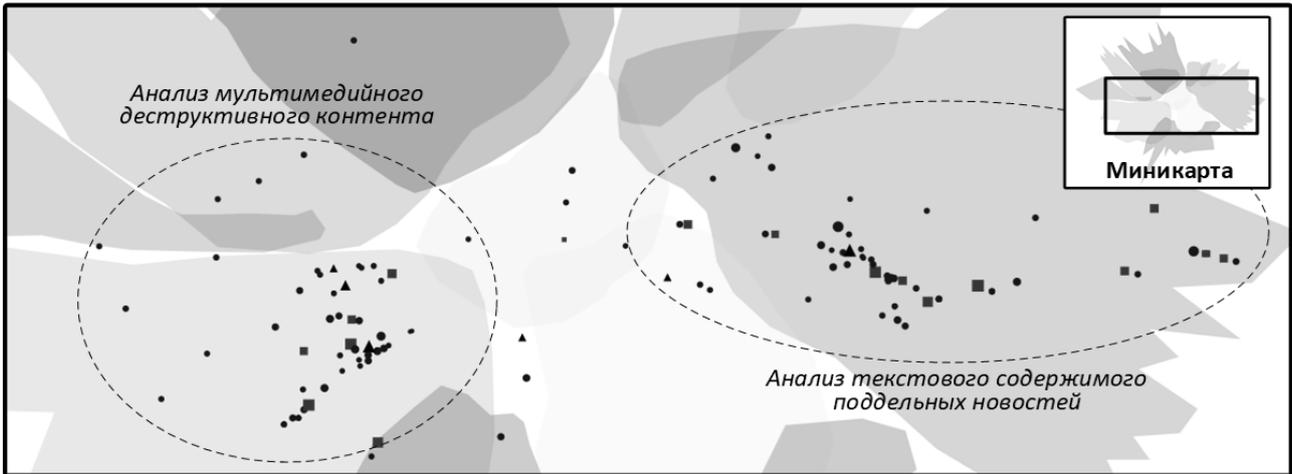


Рис. 9. Область информационной карты «Технологии искусственного интеллекта и машинного обучения (Papers With Code)», в которой расположены объекты, связанные с задачей защиты от деструктивного контента

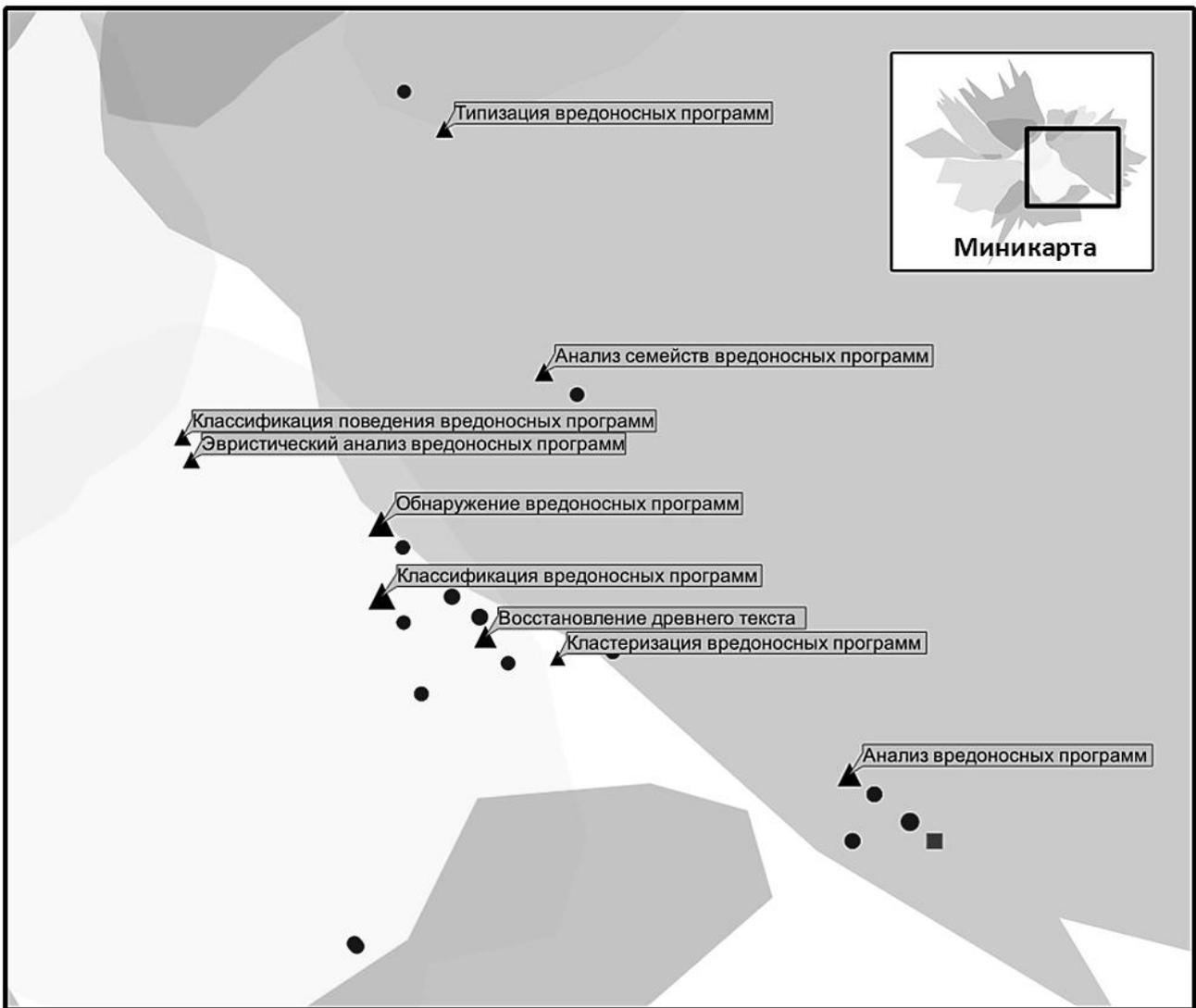


Рис. 10. Область информационной карты «Технологии искусственного интеллекта и машинного обучения (Papers With Code)», в которой расположены объекты, связанные с задачей защиты от вредоносных программ

Заключение

В данной статье были представлены результаты построения информационной карты «Технологии искусственного интеллекта и машинного обучения (Papers With Code)», на базе которой реализован блок навигации по технологиям искусственного интеллекта и машинного обучения, разрабатываемый в рамках работ по созданию киберполигона, выполняемых кафедрой систем информационной безопасности Воронежского государственного технического университета» (демонстрация работы данного блока доступна по адресу https://cybermaps.ru/projects/ai_paperswithcode/index.html [17]).

Также в статье приведены примеры решения с его помощью следующих практических задач:

- представление знаний о текущем состоянии в области технологий искусственного интеллекта и машинного обучения путём наглядного многоуровневого отображения взаимосвязей между научными публикациями, наборами данных, задачами, методами и моделями машинного обучения, а также связей между группами таких объектов;

- повышение эффективности научного поиска наборов данных и результатов исследований в области обеспечения информационной безопасности и защиты информации с помощью технологий искусственного интеллекта.

Дальнейшее развитие блока навигации по технологиям искусственного интеллекта и машинного обучения может быть осуществлено по следующим направлениям:

- расширение состава и актуализация информационных карт;

- учёт в ландшафте информационной карты дополнительных сущностей «Программный код» и «Метрики сравнения эффективности моделей машинного обучения».

Список литературы

1. Поисковая система научных

публикаций Google Академия // Информационный ресурс компании Google. URL: <https://scholar.google.com> (дата обращения: 12.05.2023).

2. Поисковая система научных публикаций ResearchGate // Информационный ресурс компании ResearchGate. URL: <https://www.researchgate.net/> (дата обращения: 12.05.2023).

3. Поисковая система научных публикаций Semantic Scholar // Информационный ресурс института The Allen Institute for Artificial Intelligence. URL: <https://www.semanticscholar.org/> (дата обращения: 12.05.2023).

4. Платформа обмена моделями машинного обучения Hugging Face // Информационный ресурс компании Hugging Face, URL: <https://huggingface.co/> (дата обращения: 12.05.2023).

5. Платформа Papers With Code // Информационный ресурс. URL: <https://paperswithcode.com/> (дата обращения: 12.05.2023).

6. Чат-бот с искусственным интеллектом ChatGPT // Информационный ресурс компании OpenAI. URL: <https://chat.openai.com/chat> (дата обращения: 12.05.2023).

7. Остапенко А.Г. Картография защищаемого киберпространства / А.Г. Остапенко, А.Л. Сердечный, А.О. Калашников; Серия Теория сетевых войн; Вып. 7. [Под ред. чл.-корр. РАН Д.А. Новикова.

8. Портал Wikidata // Информационный ресурс сообщества Wikimedia Foundation. URL: <https://wikidata.org/> (дата обращения: 12.05.2023).

9. OpenStreetMap // Информационный ресурс сообщества OpenStreetMap. URL: <https://www.openstreetmap.org/> (дата обращения: 12.05.2023).

10. Калашников А.О, Сердечный А.Л., Остапенко А.Г. Картографический подход в библиометрическом исследовании отечественных научных школ, сложившихся в области защиты информации и обеспечения информационной безопасности. // *Информация и безопасность*. 2019. Т. 22. Вып. 4. С. 455-484.
11. Сердечный А.Л. Картографическое исследование Blockchain-транзакций и смарт-контрактов киберпреступников, атакующих автоматизированные информационные системы, и оценка ущербов от реализации их атак / А.Л. Сердечный, Д.А. Скогорева, Е.П. Длинный, и др. // *Информация и безопасность*. 2021. Т. 24. Вып. 4. С. 471-500.
12. Гончаров А.А. Систематизация сведений об ошибках программного обеспечения с использованием информационной карты и оценка их значимости / А.А. Гончаров, М.А. Тарелкин, А.Л. Сердечный // *Информация и безопасность*. 2021. Т. 25. Вып. 2. С. 295-310.
13. Сердечный А.Л. Информационно-картографические системы как инструментальная основа картографии защищаемого киберпространства / Системы управления и информационные технологии. 2021. № 4 (86). С. 41-46.
14. Сердечный А.Л. К вопросу о создании платформы картографирования рисков защищаемого киберпространства / А.Л. Сердечный, А.А. Гончаров, М.А., Булычев, А.В. Коноплин, О.С. Газизянов, Р.О. Дыкин, Д.С. Нестеров, Д.А. Нархов // *Информация и безопасность*. 2021. Т. 24. Вып. 4. С. 593-600.
15. Jacomy M. ForceAtlas2, a continuous graph layout algorithm for handy network visualization designed for the Gephi software / M. Jacomy, T. Venturini, S Heymann b и др. // *PloS one*. 2014. Т. 9. №. 6. С. e98679.
16. Traag V.A. From Louvain to Leiden: guaranteeing well-connected communities / V.A. Traag, L. Waltman, N.J. Van Eck // *Scientific reports*. 2019. Т. 9. №. 1. С. 1-12.
17. Интерактивная информационная карта «Технологии искусственного интеллекта и машинного обучения (Papers With Code)» // Информационный ресурс портала CyberMaps, URL: https://cybermaps.ru/projects/ai_paperswithcode/index.html (дата обращения: 12.05.2023).

Государственный научно-исследовательский испытательный институт
проблем технической защиты информации ФСТЭК России
State science research experimental institute of technical information protection
problem of Federal service of technical an export control

Воронежский государственный технический университет
Voronezh State Technical University

Поступила в редакцию 15.05.2023

Информация об авторах

Сердечный Алексей Леонидович – канд. техн. наук, начальник лаборатории, Государственный научно-исследовательский испытательный институт проблем технической защиты информации ФСТЭК России, e-mail: alex-voronezh@mail.ru

Труфанов Артем Тимурович – студент, Воронежский государственный технический университет, e-mail: trartem2000@mail.ru

Карданов Ахмед Ануарович – студент, Воронежский государственный технический университет, e-mail: toshka0117@mail.ru

CREATING A CYBERPOLYGON: A NAVIGATION BLOCK ON TECHNOLOGIES OF ARTIFICIAL INTELLIGENCE AND MACHINE LEARNING

A.L. Serdechnyi, A.T. Trufanov, A.A. Kardanov

The article presents the results of the development of the navigation block on artificial intelligence and machine learning technologies, carried out as part of the creation of a cyber polygon. The block is based on the software implementation of the interactive information map "Artificial Intelligence and Machine Learning Technologies". The information map systematizes information about tasks, methods and models, machine learning, as well as data sets and scientific publications, which are placed on the information resource Papers With Code.

The information map reveals the complex structure of relationships between more than 3,000 machine learning problems, 34,000 scientific publications devoted to their solution, as well as 7,000 datasets, 2,000 machine learning methods and algorithms, and 8,000 models mentioned in such articles. During research of an information map it is possible to receive knowledge of a current condition of development of technologies of an artificial intellect that promotes deeper understanding of the basic tendencies in the given area, and also raises efficiency of search on the considered subject.

Also an important practical value of the developed information map is the possibility of its use for crowdsourcing work in the field of artificial intelligence. As an example the layer is built, which shows the areas of placement and relationships between tasks, models and data sets used in information security and information protection.

Navigation block is built using web-technology and is available in test mode on the portal CyberMaps.ru.

Keywords: cyber polygon, information map, information mapping, artificial intelligence, machine learning.

Submitted 15.05.2023

Information about authors

Alexey L. Serdechnyy – Cand. Sc. (Technical), Chief of Laboratory, State science research experimental institute of technical information protection problem of Federal service of technical an export control, e-mail: alexvoronezh@mail.ru

Artem T. Trufanov – student, Voronezh State Technical University, e-mail: trartem2000@mail.ru

Ahmed A. Kardanov – student, Voronezh State Technical University, e-mail: toshka0117@mail.ru