

ВІЗУАЛЬНА АНАЛІТИКА — ЕФЕКТИВНА ТЕХНОЛОГІЯ ОБРОБЛЕННЯ BIG DATA В МЕДИЦИНІ

В. В. Петров, О. П. Мінцер¹, А. А. Крючин, Є. А. Крючина²

Інститут проблем реєстрації інформації НАН України

¹*Національна медична академія післядипломної освіти імені П. Л. Шупика*

²*Київська міська клінічна лікарня № 10*

Представлено аналіз перспектив застосування візуальної аналітики в клінічній та експериментальній медицині, системі менеджменту охорони здоров'я, фармації та клінічних дослідженнях, у першу чергу для оброблення Big Data. Показано, що візуальна аналітика забезпечує більш доступний та інтуїтивно зрозумілий підхід до аналізу медико-біологічної інформації, дозволяє підвищити ефективність використання зібраних і накопичених даних, виявляти нові та невідомі знання шляхом знаходження зв'язків, патернів, трендів і аномалій у Big Data. Візуальна аналітика забезпечує управління даними, проведення їх дослідження та аналізу. Розроблені методи представлення даних у вигляді зображень, діаграм спрямовані на максимальне повне використання реєстрів медичних даних, використання накопиченої інформації для прогнозування можливості розвитку захворювань та їх профілактики та в цілому має сприяти вирішенню проблем інформаційного перевантаження. Наведені дані свідчать, що технології візуальної аналітики сприятимуть суттєвому покращенню якості медичного обслуговування населення.

Ключові слова: візуальна аналітика, Big Data, медицина.

VISUAL ANALYTICS — AN EFFECTIVE TECHNOLOGY FOR PROCESSING BIG DATA IN MEDICINE

V. V. Petrov, O. P. Mintser¹, A. A. Kryuchyn, Ye. A. Kryuchyna²

Institute for Information Recording, National Academy of Sciences of Ukraine

¹*Shupyk National Medical Academy of Postgraduate Education*

²*Kyiv City Clinical Hospital No. 10*

Background. The analysis of the prospects for the use of visual analytics in clinical and experimental medicine, a healthcare, pharmacy and clinical research, primarily for processing Big Data, is presented. The analysis shows that visual analytics provides a more accessible and intuitive approach to the analysis of biomedical information, improves the efficiency of the use of collected and accumulated data, and reveals new and unknown knowledge by finding relationships, patterns, trends and anomalies in Big Data. Visual analytics provide data management, research and analysis. The developed methods for presenting data in the form of images, diagrams are aimed at the most complete use of registers of medical data, the use of accumulated information to predict the possibility of the development of diseases and their prevention, and in general should contribute to solving the problems of information overload. The above data indicate that the technology of visual analytics will contribute to a significant improvement in the quality of medical care for the population. The purpose of the study is to analyze the current state, problems and prospects of using Visual analytics technologies for Big Data analysis in medicine.

Materials and methods. Results. Visual analytics has great potential for Big Data processing, it allows for comprehensive data analysis to improve the quality of medical care and the analysis of a large number of complex heterogeneous data of various nature. Visual analytics allows you to increase the efficiency of using the collected and accumulated data, to discover new and unknown knowledge by finding relationships, patterns, trends and anomalies in Big Data. The use of VA will allow to solve the problem of information overload, which occurs when analyzing huge data arrays, to establish a relationship between a large number of variables.

Conclusions. VA methods allow clinicians, researchers, administrators, and patients to obtain effective, meaningful information from the vast and complex data resources that a modern healthcare system offers. At the same time, there are problems with the visual study of unstructured data, the lack of standardized methods for assessing, validating and measuring the effectiveness of IA tools, which requires further research.

Key words: visual analytics, Big Data, medicine.

ВИЗУАЛЬНА АНАЛИТИКА — ЭФФЕКТИВНАЯ ТЕХНОЛОГИЯ ОБРАБОТКИ BIG DATA В МЕДИЦИНЕ

В. В. Петров, О. П. Минцер¹, А. А. Крючин, Е. А. Крючина²

Институт проблем регистрации информации НАН Украины

¹*Национальная медицинская академия последипломного образования имени П. Л. Шупика*

²*Киевская городская клиническая больница № 10*

Представлен анализ перспектив применения визуальной аналитики в клинической и экспериментальной медицине, системе менеджмента здравоохранения, фармации и клинических исследованиях, в первую очередь для обработки Big Data. Проведенный анализ показывает, что визуальная аналитика обеспечивает более доступный и интуитивно понятный подход к анализу медико-биологической информации, позволяет повысить эффективность использования собранных и накопленных данных, обнаруживать новые и неизвестные знания путем нахождения связей, паттернов, трендов и аномалий в Big Data. Визуальная аналитика обеспечивает управление данными, проведение их исследования и анализа. Разрабатываемые методы представления данных в виде изображений, диаграмм направлены на максимально полное использование реестров медицинских данных, использование накопленной информации для прогнозирования возможности развития заболеваний и их профилактики и в целом должна способствовать решению проблем информационной перегрузки. Приведенные данные свидетельствуют о том, что технологии визуальной аналитики будут способствовать существенному улучшению качества медицинского обслуживания населения.

Ключевые слова: визуальная аналитика, Big Data, медицина.

Вступлення. Сегодня объем медицинских данных в мире огромен и продолжает быстро расти в результате инновационных изменений применения принципиально новых методов анализа (статистические, контекстуальные, количественные, прогностические, когнитивные, визуальные и пр.) [1]. Отличительной чертой современной медицины является использование многоуровневых диагностических методик, систем мониторинга за пациентом, электронных медицинских карт, новых информационных технологий создания, сбора, хранения, анализа и управления медицинской информацией [2, 3]. Однако, несмотря на мнение о том, что доступ к большому количеству информации приведет к принятию более обоснованных решений, наличие всеобъемлющих и масштабных ресурсов клинических данных сделало некоторые аналитические процессы еще более трудными [4]. Big Data, вследствие особенностей структуры, эффективно используются при хранении, обработке и исследовании данных мониторинга в плохо формализуемых областях знаний, в частности, в медицине [5, 6]. Эффективное управление, анализ и интерпретация Big Data могут изменить состояние медицины, открыв новые возможности для современного здравоохранения. Аналитика Big Data в медицине и здравоохранении охватывает интеграцию и анализ большого количества сложных разнородных данных, таких как различные «омик» (геномика, эпигеномика, транскриптомика, протеомика, метаболомика, интерактомика, фармакогеномика и пр.), биомедицинских и данных электронных медицинских карт [7]. В широком смысле Big Data включают в себя структуры данных и методы для работы с ними, делящиеся на классы: структуризации, кластеризации и визуализации. Для обеспечения более глубокого подхода к анализу Big Data предлагается использовать визуальную аналитику (ВА) [5, 8]. Визуальная аналитика — продукт областей информационной и научной визуализации, которая сосредотачивается на аналитическом рассуждении, облегченном интерактивными визуальными интерфейсами [9, 10]. При эффективной организации систем поддержки принятия решений с использованием систем визуализации, многомерной структуры данных и их взаимосвязей лицо, принимающее решение, получает возможность выявления неявных связей в управляемой системе, что, в конечном итоге, на качественном уровне повышает эффективность управленческих решений [11, 12]. Визуальная

аналитика поддерживает обнаружение новых и неизвестных знаний путем нахождения связей, паттернов, трендов и аномалий в потенциально больших и сложных данных. Визуальная аналитика помогает пользователям понимать смысл разнородных данных большого объема за счет интеграции нескольких методологий анализа данных [13], обеспечивает более доступный и интуитивно понятный подход к анализу данных [8]. На сегодняшний момент разработано достаточное количество детерминистских подходов в теории распознавания образов: искусственные нейронные сети с алгоритмом обучения перцептронов; метод опорных векторов для случая линейной разделимости и линейной неразделимости классов; кластерный анализ с разными алгоритмами расстановки центров кластеров; классификация с помощью функций расстояния между векторами-образами и классами; метод потенциальных функций, позволяющий с помощью несложного алгоритма оценивать вес объектов обучающей выборки; методы дискриминантного анализа с нахождением линейной и обобщенной решающей функции [14]. В определенных случаях изображение или иконический знак, который будет рассматриваться как символ, может служить или аргументом (дополнительным аргументом) в пользу определенного решения, либо контраргументом, что приводит к фальсификации (отклонению) выдвинутой гипотезы. Таким образом, образно-лингвистическая интеллектуальная система предполагает наличие в базе знаний изображений, отражающих характерные особенности больного. Их использование может отвечать начальному состоянию процедуры формирования диагностической гипотезы или включаться в систему вывода на любом шаге. Построение такой системы может опираться на сочетание аргументов и/или правдоподобных высказываний с гештальтоподобной структурой в виде визуальных объектов ассоциативного образного ряда. Принятие диагностических решений будет осуществляться в этом случае на основе логико-аргументационных построений с включением образной компоненты [15].

Визуальная аналитика и визуализация могут использовать систему человеческого восприятия для интерпретации и выявления скрытых закономерностей в Big Data. Методы визуализации, использующие познавательные паттерны человека, применяются для выполнения анализа данных мониторинга и моделирования, отображения

иерархии классов и процессов программного обеспечения, осуществления процессов верификации и валидации, что особенно актуально для систем критического назначения [16]. Визуализация Big Data относится к областям как научной, так и информационной визуализации. В первом случае Big Data возникают в результате сложного компьютерного моделирования различных объектов и процессов. Во втором — имеет место визуальное описание и представление абстрактной информации, получаемой в результате процесса сбора и обработки многокатегориальных данных, для анализа которых необходимо применение нескольких количественных и качественных мер оценки [17]. Работы, описывающие результаты по визуализации Big Data, появились сравнительно недавно. Среди них можно выделить «White Paper» компании Intel, посвященную визуализации результатов «большого счета» [18] и «установочную» публикацию B. Shneiderman [19]. Big Data порождают предельные на данный момент случаи, что требует комплексного решения проблем компьютерной графики, визуализации, ВА и когнитивной психологии.

Анализ зарубежных работ по визуализации Big Data за последние годы показывает, что они включают в себя целый ряд направлений компьютерной визуализации, среди которых научная и информационная визуализация, визуализация программного обеспечения, визуальный анализ данных, верификация и валидация визуализации, изучение восприятия и когнитивной составляющей визуализации при использовании «больших экранов» и сред виртуальной реальности. Необходим учет этого опыта при разворачивании аналогичных исследований в Украине. Среди задач визуализации Big Data рассматриваются следующие: визуализация потоков данных; визуальный интеллектуальный анализ данных; визуальный поиск и рекомендации; описание ситуаций на основе Big Data с использованием визуализации; масштабируемые методы параллельной визуализации; современные аппаратные средства и архитектуры для анализа и визуализации данных; человеко-компьютерный интерфейс и визуализация Big Data; приложения визуализации Big Data, включая киберразведку и контрразведку, бизнес-анализ (бизнес разведку), электронную коммерцию, анализ научной информации, образование и т.д. [17].

На сегодня существенными проблемами являются ограниченные возможности человеческого познания и возрастающее давление на программные

системы со стороны данных. Основной причиной этого давления является несоответствие потребностям ВА возможностей средств аналитики и управления данными. Устранение этого несоответствия будет полезно не только для инструментов ВА, но и для аналитики и управления данными в целом. С точки зрения разработчиков программного обеспечения реализация приложения ВА затрудняется тем, что в традиционных уровнях аппаратуры и программного обеспечения отсутствуют важные сервисы, требуемые человеческим когнитивным системам при обработке сложной информации. Требования ВА не удовлетворяются даже в самых новых средствах аналитики и управления данными. Для решения этих сложных проблем требуются работы, в результате которых ВА сможет полностью справиться с поддержкой интерактивных исследований Big Data, что принципиально важно в системе здравоохранения. Инструменты ВА обычно применяются к медицинским данным, относящимся к одной из трех категорий: бизнес-проекты, клинические задачи и научные исследования в различных областях здравоохранения, в первую очередь, геномика и эпидемиология [4]. На сегодня опубликовано ряд работ применения ВА для визуализации динамических данных от множественных электронных медицинских карт, оценки эволюции симптомов во время прогрессирования заболевания, выполнения фармакокинетически-фармакодинамического анализа, построения моделей обнаружения для мониторинга заболеваний, визуализации данных о результатах лечения [20], однако проблема применения ВА в медицине как эффективной технологии обработки Big Data еще далека от своего решения.

Цель работы: анализ текущего состояния, проблем и перспектив использования технологий визуальной аналитики для анализа Big Data в медицине.

Материал и методы исследования. Проведен теоретический анализ и обобщение, систематизация результатов исследований в соответствии с рекомендациями PRISMA Group [21] с использованием таких ведущих наукометрических баз данных как ScienceDirect, PubMed, Emerald, IEEE Xplore, Taylor & Francis, печатных научных статей за период с 2009 по 2020 гг. по ключевым словам visual analytics, healthcare, Big Data, в том числе, в украинском и российском переводах.

Результаты и их обсуждение. Визуальная аналитика — наука об обнаружении знаний в больших или сложных наборах данных путем объединения

автоматизированных методов извлечения информации, машинного обучения и интеллектуального анализа данных с возможностями человеческого восприятия и когнитивных способностей [22]. Суть метода визуализации заключается в том, что исходным анализируемым данным при помощи прикладной программы визуализации ставится в соответствие некоторая их статическая или динамическая графическая интерпретация, которая визуально анализируется, а результаты анализа этой графической интерпретации (графических данных) затем истолковываются по отношению к исходным данным. Исходные данные, анализируемые методом визуализации, могут иметь различную природу. Наряду с этим, могут различаться и цели анализа исходных данных. Соответственно, могут различаться и используемые графические данные. Визуальная аналитика сочетает методы визуализации информации с методами вычислительного анализа данных, при этом отображение усиливает познавательные способности человека следующими основными способами: увеличение объема источников данных — пользователю приходится держать меньше информации в голове; сокращенный поиск — в малом пространстве может отображать значительный объем данных; поддержание связей данных, в частности, когда они размещены в пространстве по принципу их временных отношений; возможность пользователя вывести отношения или закономерности данных, которые иначе сделать было бы невозможно; слежение за большим количеством возможных вариантов событий, т. е. изменением данных; обеспечение средств динамического слежения за большим количеством параметров [23].

К прогрессивным методам визуализации, недавно получившим широкое распространение относят облако тегов, кластерграмму и исторический поток.

Облако тегов — каждому элементу в облаке тега присваивается определенный весовой коэффициент, который коррелирует с размером шрифта. В случае анализа текста величина весового коэффициента напрямую зависит от частоты употребления (цитирования) определенного слова или словосочетания. Позволяет пользователю в сжатые сроки получить представление о ключевых моментах сколько угодно большого текста или набора текстов. Кластерграмма — метод визуализации, использующийся при кластерном анализе. Показывает как отдельные элементы

множества данных соотносятся с кластерами по мере изменения их количества. Выбор оптимального количества кластеров — важная составляющая кластерного анализа. Исторический поток помогает следить за эволюцией документа, над созданием которого работает одновременно большое количество авторов. В частности, это типичная ситуация для сервисов wiki и сайта tadviser. По горизонтальной оси откладывается время, по вертикальной — вклад каждого из соавторов, т. е. объем введенного текста. Каждому уникальному автору присваивается определенный цвет на диаграмме, при этом можно провести анализ как возрастала активность авторов с течением времени [10]. Среди многообразия отдельных инструментов и систем с функциями визуализации данных Forrester Research отдельно выделяет платформы продвинутой визуализации данных. Такие платформы обладают более развитыми средствами и обеспечивают более комплексное представление информации. Вот их главные определяющие отличия:

- **динамический контент** — визуальное представление информации изменяется по мере апдейта данных в базе. Статическая визуализация, которую можно создать в большинстве офисных программ, такой возможности не предоставляет;

- **визуальные запросы** — возможность запрашивать и перезапрашивать данные путем простого манипулирования частями графики или диаграмм или специальных визуальных компонентов (выпадающих списков, вкладок и прочего);

- **мультиданные.** Типичные статические средства визуализации не могут связать и отобразить между собой более чем несколько параметров данных. В продвинутых платформах визуализации, напротив, учитываются множественные данные различных типов, прочно завязанные друг с другом, так что при обновлении какой-то части информация автоматически меняется на всех других панелях;

- **анимированная визуализация.** Некоторые параметры, например, время, могут иметь сотни и сотни значений, так что отслеживание общей картины для каждого параметра в отдельности может стать достаточно трудоемким заданием. В таких случаях на помощь приходит анимация, позволяющая запустить переход от параметра к параметру в автоматическом режиме;

- **персонализация.** То, что интуитивно понятно и необходимо одному пользователю, может совершенно не подойти другому. Кроме того, в большинстве компаний для защиты информации

и минимизации рисков предусмотрены различные уровни доступа к корпоративным данным. Платформы ADV позволяют автоматически применять персональные настройки и предпочтения;

• **использование оповещений.** Мощные ADV платформы могут подавать сигнал, когда не в состоянии корректно обработать большие массивы данных. Или, например, сигнал подается в тот момент, когда найдено какое-то критическое значение, причем оповещение может подаваться, в том числе, с помощью электронной почты и текстовых сообщений [9,10].

Визуальная аналитика становится все более популярным методом исследования и анализа Big Data в области здравоохранения [20]. Сложность медицинских процессов, индивидуальность течения заболеваний, большое количество слабо формализованных факторов и симптомов, которые подлежат анализу, с одной стороны, и наличие, с другой стороны, достаточно жёстких математических требований к объёму и согласованности к данным компьютерного анализа приводит к необходимости выявления скрытых закономерностей в статистических данных о пациентах. Решение данной задачи возможно на основе применения автоматизированных систем, которые обеспечивают врача средствами интеллектуальной поддержки и принятия решений ориентированных на обработку разнородных (количественных, качественных, текстовых) данных неограниченного объема. При этом инструментальные средства для обработки «сырых» данных должны быть просты в использовании, а результаты должны быть конкретны и понятны пользователю. Существенным этапом компьютерного анализа данных является предварительная обработка данных, включающая преобразование данных из формы, в которой их наиболее удобно собирать, в форму, требуемую для наиболее эффективного применения выбранных методов анализа [24]. Сегодня перед клиницистами стоит сложная задача анализа большого количества неструктурированных, мультимодальных данных и данных, полученных из различных источников для эффективной диагностики и мониторинга развития конкретного заболевания [25]. Информационная перегрузка — это проблема, которая возникает, когда люди пытаются проанализировать ряд переменных, которые выходят за пределы человеческого познания [9, 26]. Проблема количественной оценки человеческого потенциала при обработке информации считается

критической, которая начала активно анализироваться после классической публикации Miller G. (1956) [27]. Проведенные исследования показали, что количество переменных, которые могут быть интегрированы в одно когнитивное представление, является основным ограничением для когнитивных и нейропсихологических процессов [26]. В связи с огромным числом публикаций, исчисляемым миллионами, в частности в медицине, поиск информации превращается в серьезную проблему [28]. Анализ мультимодальных и гетерогенных данных электронных медицинских карт, зачастую содержащих также избыточные, нерелевантные и субъективные сведения, создает значительные проблемы для исследователей, пытающихся обобщить данные и получать полезные результаты для понимания сущности задачи. Существует острая необходимость в создании эффективных способов использования ВА для поддержки анализа больших объемов медицинских данных. В настоящее время разрабатываются новые интерактивные интерфейсы, чтобы создать возможность использования крупномасштабных клинических баз данных для решения широкого спектра различных задач. Клиницисты и исследователи могут использовать инструменты ВА, чтобы проводить эпидемиологический, популяционный анализ, анализ большого объема клинических, лабораторных, инструментальных данных. Администрация клиники может применять методы ВА, чтобы лучше понять продуктивность организации работы стационара, выявить проблемы в диагностике и лечении, проанализировать отзывы пациентов. Системы ВА — благодаря сочетанию передовых методов интерактивной визуализации со статистическими моделями логического вывода и корреляции — имеют потенциал для поддержки интуитивного анализа для всех этих групп пользователей [9]. Визуальные представления данных диагностики являются надежной формой анализа информации на предварительном и уточняющем этапах диагностики образного ряда визуализированных признаков. На такой основе возможна идентификация известных болезней, и выявление новых неизвестных патологий. В то же время, традиционные логико-лингвистические интеллектуальные системы не обеспечивают однозначного сопоставления между когнитивным представлением и описанием в вербальной форме визуальных характеристик заболеваний, имеющих трудно характеризуемые проявления. Причиной этого является отсутствие

в базе знаний образной компоненты [15]. Для анализа высокоинформативных изображений используются новейшие информационные технологии. Microsoft совместно с Siemens разработали новые алгоритмические возможности для анализа рентгенологических исследований. Изображения с рентгеновского аппарата отправляются для анализа одновременно в облако и врачу. В облаке есть система анализа снимков на базе искусственного интеллекта, которая обрабатывает полученное изображение и постоянно обновляется. Врач и система параллельно ставят диагноз, если есть большие расхождения, врач получает уведомление о необходимости перепроверить диагноз. Иногда бывает так, что система приняла неправильное решение, но часто она помогает увидеть скрытые вещи, которые человек не заметил [1]. Специалисты в области здравоохранения продолжают искать эффективные, экономные по времени и качественные способы обработки изображений. Разработчики баз данных ищут независимые облачные архивы для создания общего хранилища информации. Централизация архива данных даст врачам доступ ко всей необходимой информации пациента, что повысит эффективность их работы. Некоторые формы данных не соответствуют традиционной системе хранения электронной информации. Единые облачные хранилища предоставляют возможность строгой и удобной организации таких данных [29].

Визуальная аналитика и визуализация могут использовать систему человеческого восприятия для интерпретации и выявления скрытых закономерностей в Big Data, что крайне важно в онкологии и геномике, которые развиваются быстрыми темпами. Появление технологий секвенирования следующего поколения позволило создавать огромные объемы геномных данных и вызвало соответствующую потребность в новых инструментах и методах для визуализации и интерпретации этих данных. Визуализация геномных данных требует не только простого построения графика данных, но также должна предлагать решение или выбор того, что сообщение должно быть передано на конкретном графике; какие методологии следует использовать для представления результатов, они должны предоставить клиницистам, экспертам или исследователям простой, ясный и точный способ взаимодействия с данными. ВА геномных данных быстро развивается параллельно с достижениями в технологиях с высокой пропускной способностью,

таких как искусственный интеллект и виртуальная реальность [30]. Персонализированная медицина требует новых инструментов геномной визуализации, которые могут эффективно извлекать знания из геномных данных и ускорять принятие экспертных решений про оптимальное лечение каждого пациента. Однако ВА таких больших геномных данных остается серьезной проблемой. В литературе активно дискутируются инструменты, методы и тенденции для ВА геномных данных в онкологии, как традиционные подходы, в частности диаграммы рассеяния, тепловые карты, координаты и сети, а также новые технологии с использованием AI и VR [31].

Методы ВА предоставляют научную основу для эффективного контроля распространения инфекционных, в частности, респираторных заболеваний [32]. Эти исследования рассматривают распространение заболеваний, передающихся воздушно-капельным путем, как проблемы геосоциального взаимодействия, потому что человеческое взаимодействие связывает различные группы людей в географических точках с наличием вирусов, что крайне важно в настоящее время в связи с пандемией COVID 19 [33, 34]. Дополнительную ценность ВА имеет для выявления пандемий с помощью графиков взаимодействия и потоковых карт с учетом мобильности населения. Предложен визуально-аналитический подход к анализу распространения пандемии как модели пространственного перемещения, взаимодействия, контакта людей и передачи вируса. Разработанная система способна визуализировать как повседневные движения людей, так и аналогичным образом распространение заболевания в пространстве. Построенные модели пространственного взаимодействия дают ценную информацию для разработки эффективных стратегий смягчения пандемии и поддержки принятия решений в критических ситуациях [35] с учетом кластеризация заболеваемости и распространности, с разбивкой по сезонности и местоположению [25, 36]; выявление и содействие пониманию пространственно-временных горячих точек для выявления новых тенденций заболеваний и связанных с ними факторов с использованием комплексных пространственно-временных данных полученных из большого числа источников [37]. Инструменты ВА нацелены на достижение следующих взаимосвязанных задач: разработка методов визуального представления многомерных данных геосоциального взаимодействия населения

в виде геосоциальных форм, которые могут облегчить обнаружение паттернов с точки зрения распространения заболеваний и эффективности различных мер контроля, разработка методов для преобразования созданных моделей в надежные знания для поддержки процессов принятия решений в борьбе с эпидемиями и пандемиями [32]. ВА с методами искусственного интеллекта позволяют эффективно анализировать массивные данные по инфекционным заболеваниям и эпиднадзору, чтобы поддержать правительственные учреждения, поставщиков медицинских услуг и медицинских работников для реагирования на заболеваемость в будущем [38, 39].

Заключение. Осуществление процесса интерпретации результатов обработки Big Data с привлечением экспертов требует дальнейшего развития средств визуализации и отображения информации, таких как трехмерные проекции, голограммы, инструменты дополненной и виртуальной реальности [40]. Качественное повышение объема и степени восприятия данных пользователем, без отсечения существенных деталей, возможно при комплексном использовании математических методов, новых информационных технологий и учетом принципов когнитивной психологии. Ориентирование визуальных средств на сложившиеся эволюционным путем базовые познавательные структуры человеческого сознания снижает требуемый уровень минимально необходимых теоретических и специальных знаний пользователя и позволяет значительно увеличить объем одновременно воспринимаемых данных, что приводит к повышению качества принимаемых экспертных решений. Интерактивные визуальные интерфейсы ВА могут принести существенную пользу клиницистам, исследователям, пациентам и аналитикам данных. Интерактивные интерфейсы и визуальные аналитические методы могут помочь экспертам в области здравоохранения исследовать, анализировать фильтровать и передавать большие и разнообразные данные, имеющиеся в современной клинической среде [41].

Повышение уровня сложности и повсеместного распространения электронных систем

мониторинга как в медицинских учреждениях, так и в сообществе предоставляет все большие объемы медицинских данных для клиницистов, администраторов и исследователей. Методы ВА предлагают многообещающие инструменты для использования Big Data для улучшения ухода за пациентами, оценки качества лечения, управления финансами и других областей здравоохранения, хотя оценка их эффективности остается дискуссионной [42]. Визуальная аналитика может облегчить исследование Big Data, чтобы генерировать гипотезы и направлять выбор и использование расширенных методов аналитики [43]. Распространение медицинских информационных менеджментных систем и электронных медицинских карт увеличит спрос на врачей, которые могут преодолеть разрыв между медицинскими и информационными науками.

Выводы, перспективы дальнейших исследований. 1. Визуальная аналитика обладает огромным потенциалом для обработки Big Data, позволяет проводить комплексный анализ данных для улучшения качества медицинской помощи и анализа большого количества сложных гетерогенных данных различной природы.

2. Визуальная аналитика позволяет повысить эффективность использования собранных и накопленных данных, обнаруживать новые и неизвестные знания путем нахождения связей, паттернов, трендов и аномалий в Big Data.

3. Использование ВА позволит решить проблему информационной перегрузки, которая возникает при анализе огромных массивов данных, установить связь между большим числом переменных.

4. Методы ВА позволяют клиницистам, исследователям, администраторам, пациентам получить действенную, значимую информацию из обширных и сложных ресурсов данных, которые предлагает современная система здравоохранения. В то же время существуют проблемы визуального исследования неструктурированных данных, отсутствие стандартизированных методов оценки, валидации и измерения эффективности инструментов ВА, что требует дальнейших исследований.

Литература.

1. Перспективы и проблемы использования технологий Big Data в медицине / В. В. Петров, О. П. Мінцер, А. А. Крючин, Є. А. Крючина // Медична інформатика та інженерія. — 2019. — № 3 (47). — С. 20-30.
2. Проблеми зберігання медикобіологічної інформації / В. В. Петров, О. П. Мінцер, А. А. Крючин, Є. А. Крючина // Медична інформатика та інженерія. — 2017. — № 3. — С. 52-62.
3. Big Data in healthcare: management, analysis and future prospects / S. Dash, S. K. Shakyawar, M. Sharma, S. Kaushik // Journal of Big Data. — 2019. — Vol 6. — Режим доступу: <https://link.springer.com/article/10.1186/s40537-019-0217-0>.
4. Visual Analytics in Healthcare [Електронний ресурс]. — 2014. — Режим доступу: <http://visualanalyticshealthcare.org/>.
5. Шишкин Ю. Е. Визуальный анализ больших данных с применением познавательных паттернов / Ю. Е. Шишкин // Проблемы современной науки и образования. — 2017. — № 2 (84). — С. 221-225.
6. Пасынков М. А. Комплексная система интеграции баз данных мониторинга физических параметров и позиционирования в акваториях / М. А. Пасынков // Научный журнал. — 2017. — № 2 (15). — С. 29-31.
7. Ristevski B. Big Data Analytics in Medicine and Healthcare / B. Ristevski, M. Chen // J Integr Bioinform. — 2018. — Vol. 15. — № 3. — P. 2017-0030.
8. Стельмах С. VI-системы с ИИ и средствами визуальной аналитики станут основой цифрового предприятия [Електронний ресурс] / С. Стельмах. — 2018. — Режим доступу: <https://www.itweek.ru/about/authors/detail.php?ID=134156>.
9. Visual analytics in healthcare--opportunities and research challenges / Caban J. J., Gotz D. // J Am Med Inform Assoc. — 2015. — Vol. 22. — № 2. — P. 260-262.
10. Визуализация данных [Електронний ресурс]. — 2015. — Режим доступу: <http://www.tadviser.ru/index.php>.
11. Сравнительный анализ методов обнаружения изменений состояний сетевого трафика / А. В. Скатков, А. А. Брюховецкий, Ю. Е. Шишкин // Автоматизация и приборостроение: проблемы, решения: материалы междунар. науч.-техн. конфер., 2016, Севастополь: СевГУ. — С. 14-15.
12. Малярова М. В. Аналитика и визуализация «больших данных»: почему «большие данные» являются большой проблемой? / М. В. Малярова // International Scientific Review. — 2016. — № 3 (13). — С. 66-68.
13. Кузнецов С. Визуальное будущее аналитики [Електронний ресурс] / С. Кузнецов // Computer, IEEE Computer Society. — 2013. — V. 46. — № 7. — Режим доступу: <http://citforum.ru/computer/2013-07/>.
14. Трояножко О. А. Визуальная аналитика в цифровой биомедицине на примере качества диагностики рака молочной железы / О. А. Трояножко, И. Д. Колесин // International Journal of Open Information Technologies. — 2019. — Vol. 7. — № 7. — P. 27-34.
15. Кобринский Б. А. Значение визуальных образных представлений для медицинских интеллектуальных систем / Б. А. Кобринский // Искусственный интеллект и принятие решений. — 2012. — № 3. — С. 3-14.
16. Шишкин Ю. Е. Облачные сервисы в системах поддержки принятия решений / Шишкин Ю. Е. // Научный журнал. — 2017. — № 1 (14). — С. 19-20.
17. Авербух В. Л. Анализ и визуализация «больших данных» / В. Л. Авербух, Д. В. Манаков // Параллельные Вычислительные Технологии : труды междунар. науч. конф., 31.03-02.04.2015, г. Екатеринбург. — 2018. — С. 332-340.
18. Big Data Visualization: Turning Big Data into Big Insights. The Rise of Visualization-based Data Discovery Tools. White Paper. Intel IT Center [Електронний ресурс]. — 2013. — Режим доступу: <http://www.intel.com/content/dam/www/public/us/en/documents/white-papers/big-datavisualization-turning-big-data-into-big-insights.pdf>.
19. Shneiderman B. The big picture for big data: Visualization / B. Shneiderman // Science. — 2014. — Vol. 343. — № 6172. — P. 730.
20. Big data and visual analytics in anaesthesia and health care / A. F. Simpaio, L. M. Ahumada, M. A. Rehman // British Journal of Anaesthesia. — 2015. — Vol. 115. — № 3. — P. 350-356.
21. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement / D. Moher, A. Liberati, J. Tetzlaff, D. G. Altman // PLoS Med. — 2009. — Vol. 6. — № 7. — P. e1000097.
22. Yaeli A. Visual Analytics Kit for Healthcare - overview. Currently real world data (RWD) is playing an increasing role in health care decision making [Електронний ресурс] / A. Yaeli. — Режим доступу: https://researcher.watson.ibm.com/researcher/view_group.php.
23. Кашницкий Ю. С. Визуальная аналитика в задаче трикластеризации многомерных данных / Ю. С. Кашницкий // Труды МФТИ. — 2014. — Том 6. — № 3. — С. 43-56.
24. Комплексный анализ медицинских данных / Н. Э. Самойленко, В. Н. Кувина, С. С. Кувин // Вестник воронежского государственного технического университета. — 2009. — Том 5. — № 9. — С. 114-118.
25. Visual Analytic Tools and Techniques in Population Health and Health Services Research: Protocol for a Scoping Review / J. A. Chishtie, J. Babineau, I. A. Bielska et al. // JMIR Res Protoc. — 2019. — Vol. 8. — № 10. — P. e14019.
26. How many variables can humans process? / G. S. Halford, R. Baker, J. E. McCredden, J. D. Bain // Psychol Sci. — 2005. — Vol. 16. — № 1. — P. 70-76.
27. Miller G. A. The magical number seven, plus or minus two / G. A. Miller // Psychological Review. — 2018. — Vol. 63. — № 2. — P. 81-97.

28. Еляков А. Д. Дефицит и избыток информации в современном социуме [Электронный ресурс] / А. Д. Еляков. — 2010. — Режим доступа: http://ecsocman.hse.ru/data/2011/03/11/1214896871/Elyakov_11.pdf.
29. Визуализация в медицине [Электронный ресурс]. — Режим доступа: <http://zdrav.expert/index.php/file:G:/VisualAnalyticinHealthcare/.html>.
30. Benke K. Artificial Intelligence and Big Data in Public Health / K. Benke, G. Benke // *Int J Environ Res Public Health*. — 2018. — Vol. 15. — № 12. — P. pii: E2796.
31. Visual Analytics of Genomic and Cancer Data: A Systematic Review / Z. Qu, C. W. Lau, Q. V. Nguyen et al. // *Cancer Inform.* — 2019. — Vol. 18. — № 3. — P. 1176935119835546.
32. Luo W. Visual analytics of geo-social interaction patterns for epidemic control / W. Luo // *Int J Health Geogr.* — 2016. — Vol. 15. — P. 28.
33. Qiu H. J. Using the big data of internet to understand coronavirus disease 2019's symptom characteristics: a big data study / H. J. Qiu, L. X. Yuan, X. K. Huang et al. // *Zhonghua Er Bi Yan Hou Tou Jing Wai Ke Za Zhi*. — 2019. — Vol. 55. — P. E004.
34. Digital technology and COVID-19 / D. S. W. Ting, L. Carin, V. Dzau et al. // *Nat Med*. — 2020. — Vol. 26. — № 4. — P. 459-461.
35. Guo D. Visual analytics of spatial interaction patterns for pandemic decision support / D. Guo // *Int J Geogr Inf Sci*. — 2007. — Vol. 21. — № 8. — P. 859-877.
36. Dynamic maps: a visual-analytic methodology for exploring spatio-temporal disease patterns / D. A. Castronovo, K. K. Chui, E. N. Naumova // *Environ Health*. — 2009. — Vol. 8. — P. 61.
37. Visual analytics approach to understanding spatiotemporal hotspots / R. Maciejewski, S. Rudolph, R. Hafen et al. // *IEEE Trans Vis Comput Graph*. — 2010. — Vol. 16. — № 2. — P. 205-220.
38. Artificial Intelligence for infectious disease Big Data Analytics / Z. S. Y. Wong, J. Zhou, Q. Zhang // *Infect Dis Health*. — 2019. — Vol. 24. — № 1. — P. 44-48.
39. Evaluating visual analytics for health informatics applications: a systematic review from the American Medical Informatics Association Visual Analytics Working Group Task Force on Evaluation / D. T. Y. Wu, A. T. Chen, J. D. et al. // *J Am Med Inform Assoc*. — 2019. — Vol. 26. — № 4. — P. 314-323.
40. Visualizing the knowledge structure and evolution of big data research in healthcare informatics / D. Gu, J. Li, X. Li, C. Liang // *Int J Med Inform.* — 2017. — Vol. 98. — № 4. — P. 22-32.
41. The challenge of big data in public health: an opportunity for visual analytics / O. Ola, K. Sedig // *Online J Public Health Inform.* — 2014. — Vol. 5. — № 3. — P. 223.
42. Concurrence of big data analytics and healthcare: A systematic review / N. Mehta, A. Pandit // *Int J Med Inform.* — 2018. — Vol. 114. — P. 57-65.
43. A Systematic Review on Healthcare Analytics: Application and Theoretical Perspective of Data Mining / M. S. Islam, M. M. Hasan, X. Wang // *Healthcare (Basel)*. — 2018. — Vol. 6. — № 2. — P. 54.

References.

1. Petrov, V. V., Mintser, O. P., Kryuchyn, A. A., Kryuchyna, E. A. (2019). Prospects and problems of the use of technology in medicine. *Medical Informatics and Engineering*, 3 (47), 20-30. doi: <https://doi.org/10.11603/mie.1996-19602019.3.10429>. [In Russian].
2. Petrov, V. V., Mintser, O. P., Kryuchyn, A. A., Kryuchyna, E. A. (2017). Problems of medical medical information. *Medical Informatics and Engineering*, 3, 52-62. doi: <http://dx.doi.org/10.11603/mie.1996-1960.2017.3.8182>. [In Russian].
3. Dash, S., Shakyawar, S. K., Sharma, M., Kaushik, S. (2019). Big Data in healthcare: management, analysis and future prospects. *Journal of Big Data*, 6, 542.
4. Visual Analytics in Healthcare. (2014). Retrived from: <http://visualanalyticshealthcare.org/>.
5. Shishkin, Yu. E. (2017). Visual analysis of big data using cognitive patterns. *Problems of Modern Science and Education*, 2 (84), 221-225. [In Russian].
6. Pasyukov, M. A. (2017). Integrated database integration system for monitoring physical parameters and positioning in water areas. *Scientific journal*, 2 (15), 29-31. [In Russian].
7. Risteviski, B., Chen, M. (2018). Big Data Analytics in Medicine and Healthcare. *J Integr Bioinform*, 15 (3), 2017-0030. doi: 10.1515/jib-2017-0030.
8. Stelmakh, S. (2018). VA-systems with AI and visual analytics will become the basis of a digital enterprise. Retrived from <https://www.itweek.ru/about/authors/detail.php?ID=134156>. [In Russian].
9. Caban, J. J., Gotz, D. (2015). Visual analytics in healthcare-opportunities and research challenges. *J Am Med Inform Assoc.*, 22 (2), 260-262. doi: 10.1093/jamia/ocv006.
10. Data visualization. (2015). Retrived from: <http://www.tadviser.ru/index.php>.
11. Skatkov, A. V., Bryukhovetsky, A. A., Shishkin, Yu. E. (2016). Comparative analysis of methods for detecting changes in network traffic states. *Automation and Instrument-Making: Problems, Solutions: International Materials. scientific and technical confer. Sevastopol, SevSU*, 14-15. [In Russian].
12. Malyarova, M. V. (2016). Analytics and visualization of «big data»: why is «big data» a big problem? *International Scientific Review*, 3 (13), 66-68. [In Russian].
13. Kuznetsov, S. (2013). The Visual Future of Analytics. *Review of Computer Journal (IEEE Computer Society, V. 46, No. 7, July 2013)*. Retrived from: <http://citforum.ru/computer/2013-07/>. [In Russian].

14. Troyanozhko, O. A., Kolesin, I. D. (2019). Visual analytics in digital biomedicine as an example of the quality of diagnosis of breast cancer. *International Journal of Open Information Technologies*, 7 (7), 27-34. [In Russian].
15. Kobrinsky, B. A. (2012). The Importance of Visual Imagery for Medical Intelligent Systems. *Artificial Intelligence and Decision Making*, 3, 3-14. [In Russian].
16. Shishkin, Yu. E. (2017). Cloud services in decision support system. *Scientific journal*, 1 (14), 19-20. [In Russian].
17. Averbukh, V. L., Manakov, D. V. (2018). Analysis and visualization of «big data». Proceedings of the international scientific conference «Parallel Computing Technologies» (PaVT'2015). *Yekaterinburg, March 31 - April 2, 2015*, 332-340. [In Russian].
18. Big Data Visualization: Turning Big Data into Big Insights. The Rise of Visualization-based Data Discovery Tools. White Paper. *Intel IT Center (2013)*. Retrived from: <https://www.intel.com/content/dam/www/public/us/en/documents/white-papers/big-data-visualization-turning-big-data-into-big-insights.pdf>.
19. Shneiderman, B. (2014). The big picture for big data: Visualization. *Science*, 343 (6172), 730. doi: 10.1126/science.343.6172.730-a.
20. Simpao, A. F., Ahumada, L. M., Rehman, M. A. (2015). Big data and visual analytics in anaesthesia and health care. *Br J Anaesth.*, 115(3), 350-6. doi: 10.1093/bja/aeu552.
21. Moher, D., Liberati, A., Tetzlaff, J., Altman, D. G. (2009). Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *PLoS Med.*, 6 (7), e1000097. doi: 10.1371/journal.pmed.1000097.
22. Yaeli, A. (2020). Visual Analytics Kit for Healthcare - overview. Currently real world data (RWD) is playing an increasing role in health care decision making. Retrived from: https://researcher.watson.ibm.com/researcher/view_group.php.
23. Kashnitsky, U. S. (2014). Visual Analytics in the Triclusterization. Problem multidimensional data Proceedings of MIPT, 6 (3), 43-56. [In Russian].
24. Samoilenko, N. E., Kuvina, V. N., Kuvina, S. S. (2009). Comprehensive analysis of medical data. Bulletin of the Voronezh State Technical University, 5 (9), 114-118. [In Russian].
25. Chishtie, J. A., Babineau, J., Bielska, I. A. et al. (2019). Visual Analytic Tools and Techniques in Population Health and Health Services Research: Protocol for a Scoping Review. *JMIR Res Protoc.*, 8 (10), e14019. <http://doi.org/10.2196/14019>.
26. Halford, G. S., Baker, R., McCredden, J. E., Bain, J. D. (2005). How many variables can humans process? *Psychol Sci.*, 16 (1), 70-6. doi.org/10.1111%2Fj.0956-7976.2005.00782.x.
27. Miller, G. A. (1956). The magical number seven, plus or minus two: some limits on our capacity for processing information. *Psychological Review*, 63 (2), 81-97.
28. Elyakov, A. D. (2010). Deficit and excess of information in modern society. Retrived from http://ecsocman.hse.ru/data/2011/03/11/1214896871/Elyakov_11.pdf. [In Russian].
29. Medical Imaging. Retrived from: <http://zdrav.expert/index.php/file:G:/VisualAnalyticinHealthcare/.html>.
30. Benke, K., Benke, G. (2018). Artificial Intelligence and Big Data in Public Health. *Int J Environ Res Public Health*. Dec, 10; 15 (12), pii: E2796. doi: 10.3390/ijerph15122796.
31. Qu, Z, Lau, C. W, Nguyen, Q. V., Zhou, Y., Catchpoole, D. R. (2019). Visual Analytics of Genomic and Cancer Data: A Systematic Review. *Cancer Inform.*, 18:1176935119835546. doi: 10.1177/1176935119835546.
32. Luo W. (2016). Visual analytics of geo-social interaction patterns for epidemic control. *Int J Health Geogr.*, 15, 28. doi: 10.1186/s12942-016-0059-3.
33. Qiu, H. J., Yuan, L. X., Huang, X. K., Zhou, Y. Q. et al. (2020). Using the big data of internet to understand coronavirus disease 2019's symptom characteristics: a big data study. *Zhonghua Er Bi Yan Hou Tou Jing Wai Ke Za Zhi*. 55 (0), E004. doi: 10.3760/cma.j.cn115330-20200225-00128.
34. Ting, D. S. W., Carin, L., Dzau, V., Wong, T. Y. (2020). Digital technology and COVID-19. *Nat Med*. Apr., 26 (4), 459-461. doi: 10.1038/s41591-020-0824-5.
35. Guo D. (2007). Visual analytics of spatial interaction patterns for pandemic decision support. *Int J Geogr Inf Sci.*, 21 (8), 859-77. doi: 10.1080/13658810701349037.
36. Castronovo, D. A., Chui, K. K., Naumova, E. N. (2009). Dynamic maps: a visual-analytic methodology for exploring spatio-temporal disease patterns. *Environ Health.*, 8, 61. doi: 10.1186/1476-069X-8-61.
37. Maciejewski, R., Rudolph, S., Hafen, R. et al. (2010). A visual analytics approach to understanding spatiotemporal hotspots. *IEEE Trans Vis Comput Graph.* 16 (2), 205-220. doi: 10.1109/TVCG.2009.100. *Int J Health Geogr*. 2016; 15: 28.
38. Wong Z. S. Y. (2-19). Artificial Intelligence for infectious disease Big Data Analytics. *Infect Dis Health*, 24:1, 44-48. doi: 10.1016/j.idh.2018.10.002.
39. Wu, D. T. Y., Chen, A.T., Manning, J. D. et al. (2019). Evaluating visual analytics for health informatics applications: a systematic review from the American Medical Informatics Association Visual Analytics Working Group Task Force on Evaluation. *J Am Med Inform Assoc.*, 26 (4), 314-323. doi: 10.1093/jamia/ocy190.
40. Gu, D., Li, J., Li, X., Liang, C. (2017). Visualizing the knowledge structure and evolution of big data research in healthcare informatics. *Int J Med Inform.*, 98, 22-32. doi: 10.1016/j.ijmedinf.2016.11.006.

41. Ola, O., Sedig, K. (2014). The challenge of big data in public health: an opportunity for visual analytics. Online *J Public Health Inform.*, 5 (3), 223. doi: 10.5210/objphi.v5i3.4933. eCollection 2014.
42. Mehta, N., Pandit, A. (2018). Concurrence of big data analytics and healthcare: A systematic review. *Int J Med Inform.*, 114, 57-65. doi: 10.1016/j.ijmedinf.2018.03.013.
43. Islam, M. S., Hasan, M. M., Wang, X., Germack, H. D., Noor-E-Alam, M. (2018). A Systematic Review on Healthcare Analytics: Application and Theoretical Perspective of Data Mining. *Healthcare (Basel)*, 6 (2), 54. doi:10.3390/healthcare6020054.