

## Количественная оценка когнитивной интерпретируемости визуализации

А.А. Захарова<sup>1\*</sup>, Е.В. Вехтер<sup>2</sup>, А.В. Шкляр<sup>3</sup>, А.В. Крысько<sup>4</sup>, О.А. Салтыкова<sup>5</sup>

\* Институт прикладной математики им. М.В. Келдыша РАН

<sup>1,2,3</sup> Томский политехнический университет, Томск, Россия

<sup>4,5</sup> Кафедра "Математика и моделирование" СГТУ, Саратов, Россия

<sup>1</sup> ORCID: 0000-0003-4221-7710, [zaa@tpu.ru](mailto:zaa@tpu.ru)

<sup>2</sup> ORCID: 0000-0003-0604-0399, [vehter@tpu.ru](mailto:vehter@tpu.ru)

<sup>3</sup> ORCID: 0000-0003-4442-7420, [shklyarav@tpu.ru](mailto:shklyarav@tpu.ru)

<sup>4</sup> ORCID: 0000-0002-9389-5602, [anton.krysko@gmail.com](mailto:anton.krysko@gmail.com)

<sup>5</sup> ORCID: 0000-0002-3880-6662, [olga\\_a\\_saltykova@mail.ru](mailto:olga_a_saltykova@mail.ru)

### Аннотация

В работе показаны результаты исследований когнитивной интерпретируемости визуализации и предложены подходы к получению эмпирических оценок практической результативности средств визуализации. Полученные результаты использованы при разработке технологии когнитивной интерпретации гетерогенных данных и необходимых для этого средств визуальной аналитики.

**Ключевые слова:** визуальная аналитика, визуальная модель, анализ данных, интерпретация, визуальное восприятие.

## 1. Введение

Специфичным эффектом использования визуализации, согласно [10], является влияние на сознание наблюдателя через формирование новой информационной реальности. Объяснение этого результата основано на изучении влияния иллюзии объективности визуализации на результат ее интерпретации. Для систем научной визуализации использование и учет всех особенностей когнитивной интерпретируемости визуализации является одной из первоочередных задач [2, 7], решение которой затруднено отсутствием систематизированного подхода к визуальной аналитике.

Исследователями [3] приведен рейтинг нерешенных проблем визуализации. К настоящему времени, этот список может быть уточнен с учетом актуальных задач, к решению которых может

быть привлечена визуализация, в т.ч. научная и когнитивная:

1. Количественная мера. Проблема сравнения и выбора средств визуализации из-за отсутствия обобщенной схемы визуализации и соответствующей ей системы критериев оценки.
2. Юзабилити. Проблема достижения средствами визуализации пользовательских характеристик, обеспечивающих возможность получения достоверного решения поставленной задачи исследования при использовании лишь визуальной интерпретации.
3. Перцепция. Задача увеличения информативности и интерпретируемости средств визуальной аналитики в результате направленного использования потенциала визуального восприятия.
4. Предварительная информированность. Проблема избыточного информирования пользователя, сни-

жающего когнитивную результативность визуализации и увеличивающего ресурсоемкость средств визуализации.

5. Обучение. Увеличение сложности задач визуального исследования и создаваемых средств визуальной аналитики создает необходимость предварительного обучения пользователей.
6. Масштабируемость. Задача поиска способов визуализации информации, сохраняющих высокую когнитивную интерпретируемость при любых изменениях детализации или объема исследуемых данных.
7. Интерпретируемая эстетика. Исключение или использование субъективного критерия визуальной эстетики для достижения цели исследования.
8. Динамика. Задача использования восприятия изменяющихся образов для усиления когнитивной интерпретируемости визуализации.
9. Внутрисистемная логика. Формализация использования визуализации в качестве знаковой языковой системы для получения новых знаний при использовании лишь внутрисистемных операций.
10. Представление знаний. Решение комплексной задачи эффективного хранения, передачи, получения и использования информации в резуль-

тате создания ее визуального представления.

Когнитивная интерпретация визуальных образов предъявляет к разработчикам средств визуализации требования, выполнение которых затруднено необходимостью направленного использования особенностей восприятия и мышления, многие из которых недостаточно изучены [1, 4]. В работе представлены результаты, полученные в результате исследования особенностей визуального восприятия и когнитивной интерпретации абстрактных образов.

## 2. Балансовая модель оценки визуализации

В работе [8] введено понятие структурной единицы визуализации, позволившее систематизировать усилия разработчиков средств визуализации, предназначенных для решения задач различного типа и уровня сложности. Объединение структурных единиц в функционирующую систему визуальной аналитики, приводит к построению визуальной модели, когнитивное значение которой превосходит результативность исследования отдельных образов. Эмерджентность визуальной модели обеспечивается связями составляющих ее структурных элементов.

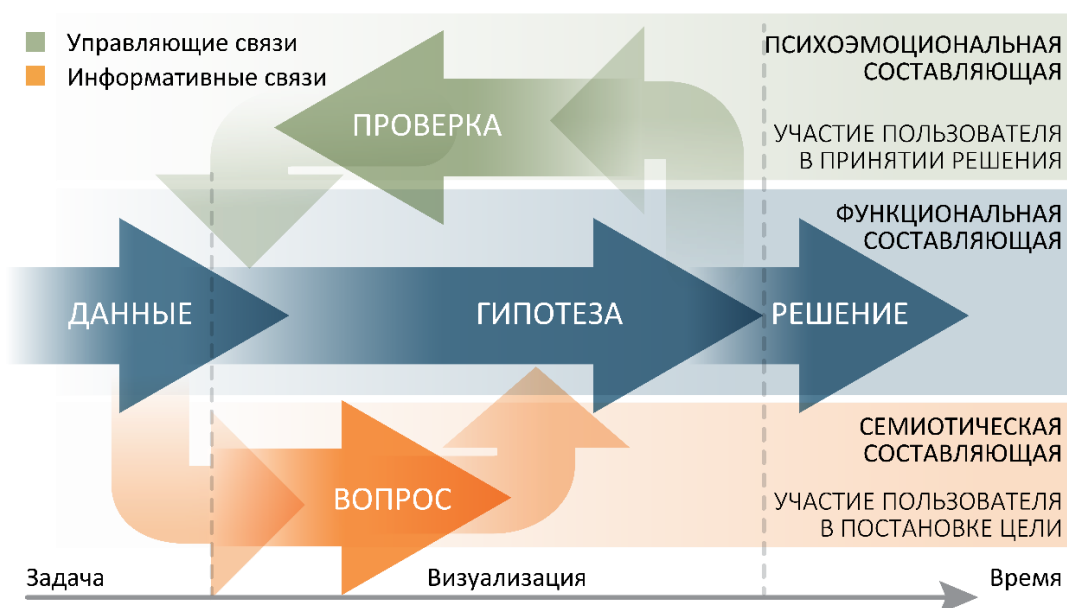


Рис. 1. Информативные и управляющие связи в процессе визуализации.

На основании различий в способах конечной реализации и практического использования структурных элементов визуальной модели, связи между ними разделены на две функциональные группы: информативные и управляющие связи (

Рис. 1). Информативные связи — процессы передачи необходимых данных между структурными единицами, создающие основу для построения образа данных, интерпретируемого при решении поставленной задачи. Управляющие связи — процессы управления использованием данных. Управление включает в себя принятие решения о форме визуального образа, необходимой для ответа на вопрос, а также о соответствии между вопросом и сформулированной гипотезой.

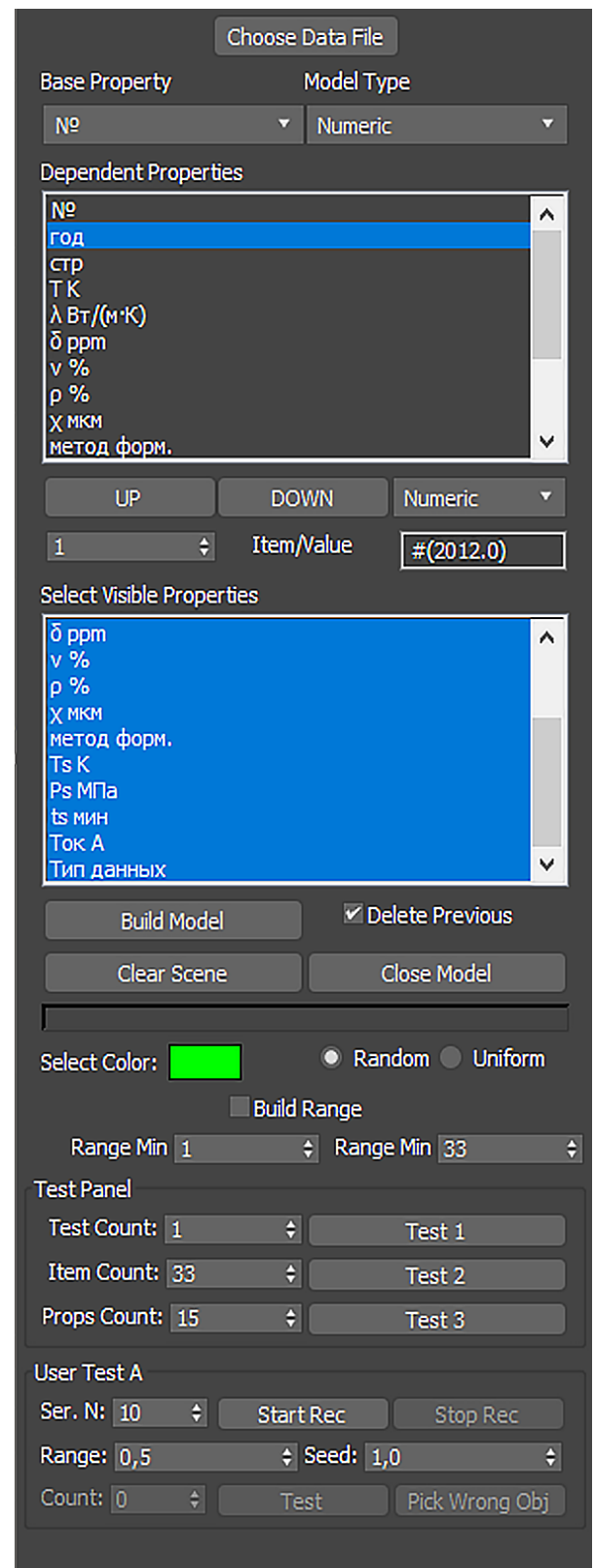


Рис. 2. Интерфейс ПО «Анализатор визуального представления».

На основании результатов решения ряда практических задач, введена балансовая модель общей результативности визуализации  $W_F$ , объединяющая позитивные и негативные факторы, характеризующие процессы визуализации, в том числе когнитивную интерпретируемость. Таким образом, общая результативность:

$$W_F = W^+(K_{TRG}(T_A)) + W^+(K_{ADD}(T_A)) - W^C(T_A) - W^M(T_A),$$

где  $W^+(K_{TRG}(T_A))$  – ценность решения задачи анализа,  $W^+(K_{ADD}(T_A))$  – ценность дополнительных знаний, полученных в результате формулирования и проверки гипотез решения,  $W^C(T_A)$  – ценность ресурсов, израсходованных для создания визуальной модели,  $W^M(T_A)$  – ценность ресурсов, использованных для управления свойствами модели в процессе получения решения.

### 3. Программное обеспечение

Для получения эмпирических данных, подтверждающих разработанную схему визуализации, а также для определения факторов, влияющих на ее результативность, предложено и разработано соответствующее ПО «Анализатор визуального представления» (Рис. 2). В основе алгоритма данного ПО лежит утверждение о прямой зависимости между результативностью процесса визуального исследования и временем, затрачиваемым пользователем на достижение цели анализа. Таким образом, основной задачей для разработчиков средств визуализации данных следует считать сокращение времени визуального исследования. Использование ПО «Анализатор визуального представления» позволяет получать данные о зависимости времени исследования на каждом этапе решения задачи анализа

от любых факторов, оказывающих на него существенное влияние.

### 4. Методика исследования

Получение и корректная интерпретация эмпирических данных создает условия для построения средств визуализации, обладающих обоснованным и направляемым когнитивным эффектом [5]. Для достижения этой цели разработана методика исследования возможностей визуального представления, апробированная на нескольких вариантах средств визуального анализа многомерных гетерогенных данных. Методика состоит в проведении серий решений тестовых задач, сопровождаемых рядом контролируемых ограничений, и измерении интервалов времени взаимодействия исследователя и визуальной модели. Взаимодействие пользователя с моделью, в этом случае, подразумевает любые доступные пользователю операции, за исключением изменения функции визуального представления. Допустимыми операциями является адаптация визуального представления данных к особенностям восприятия пользователя, фильтрация данных, переход к следующей задаче, чтение атрибутивных значений визуализированных данных и некоторые другие.

Тестовое решение предполагает создание образа исследуемых данных с использованием predeterminedного способа визуального представления (Рис. 3). Выбор первоначального способа визуализации соответствует типу задачи исследования (задача обучения, информирования, принятия решения) и не зависит от предпочтений исследователя. Таким образом, в начальный момент времени пользователь располагает визуальной моделью данных и знает цель (вопрос) решаемой им задачи [9].

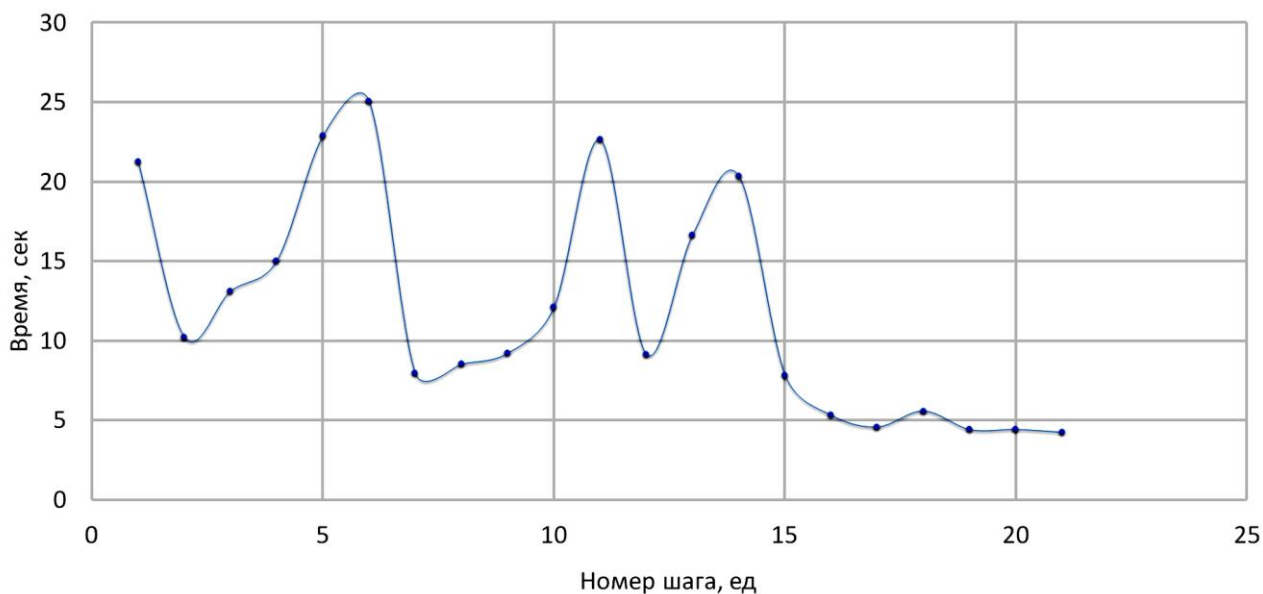


Рис. 3. Изменение скорости интерпретации визуализации.

В ходе тестового решения, исследователю, анализирующему образ данных, позволяется формулировать неограниченное число гипотез ответа на вопрос задачи, каждая гипотеза считается шагом анализа. Формулирование верной гипотезы означает завершение тестового решения. Цикличность процедуры решения позволяет автоматически фиксировать продолжительность каждого этапа для последующего анализа результативности визуализации.

Полученные результаты позволили сделать ряд заключений о характере взаимодействия пользователя и визуальной модели в процессе решения задачи анализа. Прежде всего, предложенная методика изучения когнитивной интерпретируемости визуализации позволила оценить общее время и влияние ряда факторов на длительность решения. Например, в случае проведения серии измерений с участием одного пользователя это позволило оценить индивидуальные особенности когнитивных процессов. Типичный пример результатов таких измерений (Рис. 3) показывает существенное изменение скорости мышления исследователя в ходе решения задач исследования любого заданного типа.

## 5. Результаты

По результатам проведенных измерений, в процессе взаимодействия пользователя и визуальной модели выделены интервалы, на протяжении которых время, затрачиваемое пользователем на построение очередной гипотезы, непрерывно сокращается. Во многих случаях, интервалы, обладающие подобными признаками, расположены в начале процесса анализа, однако, могут возникать и на более поздних шагах решения. В соответствии с определением структурной единицы визуализации и определением аналитической визуальной модели, подобные интервалы, возникающие на начальных шагах (Рис. 4), могут быть определены как этапы обучения, соответствующие ознакомлению пользователя с использованной функцией визуального представления. Выделение подобного этапа ознакомления позволяет ввести качественную оценку функции визуального представления, определяющую согласованность восприятия конкретного пользователя и свойств выбранного способа визуализации данных.

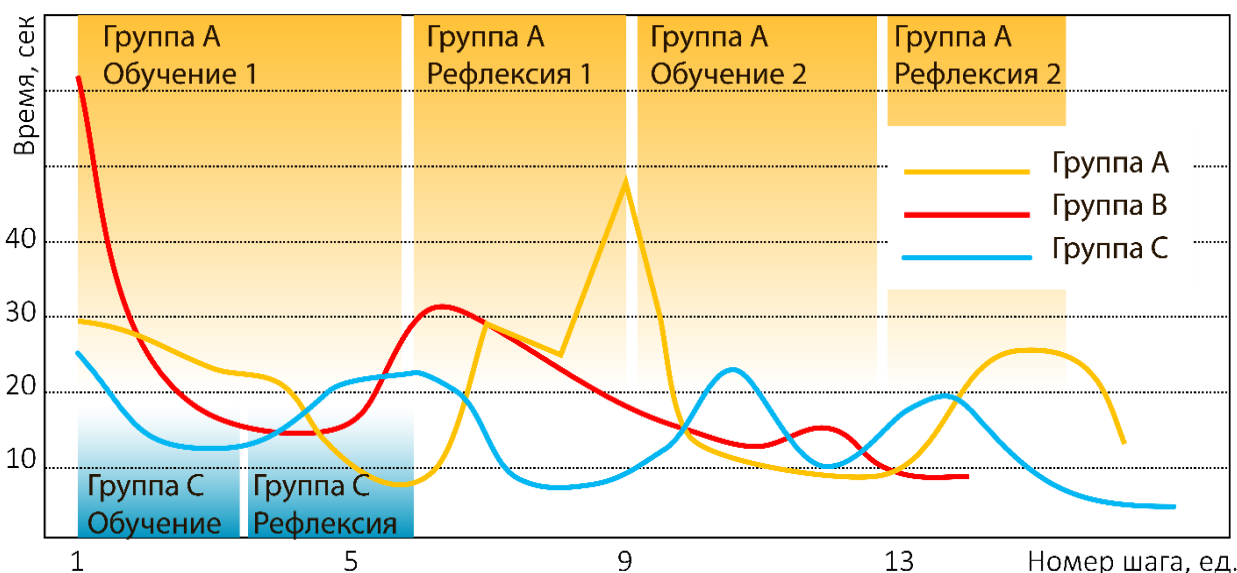


Рис. 4. Периоды обучения и рефлексии.

На основании проведенных измерений, показано влияние индивидуальных свойств исследователя на продолжительность интервала обучения, завершающегося переходом к быстрому построению новых гипотез. К индивидуальным особенностям восприятия, по результатам проведенных измерений, отнесены как возможности визуального восприятия, так и предварительная информированность пользователя, влияющая на итоги когнитивной интерпретации визуализации. Направленное использование предварительной информированности пользователя и свойств его восприятия являются способами увеличения результативности визуального анализа.

Изменение скорости построения новых гипотез, наблюдаемое на протяжении процесса анализа, носит периодический характер. В процессе решения задачи анализа пользователь не получал сведений из внешних источников, поэтому изменение времени, затрачиваемого на каждом шаге анализа, может быть объяснено изменением соотношения между прямыми и обратными процессами [6]. Незначительное снижение скорости решения свидетельствует об увеличении времени, затрачиваемом на построение и проверку очередной гипотезы, и может интерпретироваться как сомнение в правильности совершаемых

действий, вызванное ошибками, допущенными на предыдущих шагах. Значительное снижение скорости анализа, возникающее с периодичностью, различающейся у разных пользователей, может интерпретироваться как потребность пользователя в дополнительной паузе перед построением новой гипотезы. В большинстве измерений, резкое увеличение времени построения новой гипотезы возникает после серии быстрых ошибочных предположений, что говорит об отсутствии у пользователя понимания информации, представленной в визуальном образе.

На основании сделанных наблюдений за ходом решения задач анализа, проводимого различными пользователями, возможно предположение, объясняющее возникающее замедление необходимостью переосмысления допущенных ошибок. Завершение этапа рефлексии соответствует увеличению информированности исследователя, а его длительность также зависит от индивидуальных особенностей исследователя.

## 6. Оценка интерпретируемости

Построение схемы процесса визуального исследования, использующей понятие структурной единицы визуализации, сделало возможным разработку

Таблица 1.  
Результаты тестовых решений

Построение модели, сек	Эфф-ть классификации, %	Средняя эфф-ть, %
150,34		
190,12	26,46	26,46
201,23	33,85	30,15
202,16	34,47	31,59
200,43	33,32	32,02
191,45	27,34	31,09
192,51	28,05	30,58
178,10	18,46	28,85
193,16	28,48	28,80
164,07	9,13	26,62
199,21	32,51	25,21
169,52	12,76	24,17
180,28	19,91	23,84
178,99	19,06	23,50
183,43	22,01	23,40
161,30	7,29	22,40
191,01	27,05	22,67
199,43	32,65	23,22
156,32	3,98	22,21
185,53	23,41	21,43
193,25	28,54	21,71
192,21	27,85	21,95

ряда понятий и процессов, необходимую для ответа на общие вопросы визуализации, указанные выше. В числе таких новых результатов следует упомянуть алгоритм построения средств визуализации, соответствующих требованиям конкретной решаемой задачи, методику проведения визуального исследования, направленную на увеличение его результативности, а также общую классификацию задач визуализации. Предложенная методика, использующая ПО «Анализатор визуального представления», позволяет получить численные оценки изменения результативности визуализации, возникающего благодаря привлечению этих результатов к решению практических задач исследования данных.

Создание классификации задач визуализации дало несколько положительных результатов, с точки зрения когнитивной интерпретируемости визуализации. Во-первых, происходит упрощение и формализация части процессов, необходимых для создания новых средств визуализации. Во-вторых, наличие классификации задач визуализации делает возможным накопление и систематизацию опыта пользователем. Увеличение предварительной информированности пользователя, в свою очередь, обеспечивает условия для успешной когнитивной интерпретации визуализации при решении новых задач.

Разработанная классификация [8] позволяет увеличить результативность создания или выбора средств визуальной аналитики, необходимых для обработки гетерогенных данных, благодаря обоснованию требований, предъявляемых к средствам визуализации (около 25% сокращение времени выбора средства визуальной аналитики). Для оценки результативности использованы данные (**Ошибка! Источник ссылки не найден.**), полученные в ходе тестовых измерений. Предполагается, что использование классификации позволяет определить тип задачи визуализации и безошибочно выполнить алгоритм построения средства визуального исследования, используя данные пред-

варительного исследования. Увеличение результативности в случае увеличения объема исходных данных, проявляющееся в сокращении времени анализа, является основанием для использования средств визуальной аналитики в качестве инструментов обработки эмпирических данных для слабо формализованных задач.

При определении факторов, оказывающих заметное влияние на когнитивную интерпретацию визуализации, особое внимание было уделено результативности комплексного подхода к визуализации [11]. Предполагается, что применение комплексного подхода сокращает время использования структурной единицы визуализации в следствие изменения субъективной информированности пользователя. Для получения численной оценки результативности комплексного подхода предложено сравнение длительностей суммарного интервала обучения и рефлексии для последовательности решений однотипных задач. Полученные экспериментальные данные позволили сделать вывод о том, что предложенный комплексный подход к визуализации и интерпретации гетерогенных данных позволяет увеличить результативность (более 40% сокращение времени формирования гипотезы) средств визуальной аналитики в результате совместного использования вычислительных и когнитивных ресурсов.

## 7. Заключение

Предложены и разработаны средства получения экспериментальных оценок результативности средств визуального анализа. Полученные значения позволили определить степень влияния факторов, входящих в общее определение результативности визуального анализа, на ее значение

Получены сравнительные значения результативности этапов создания средств визуального анализа, использующих предложенный алгоритм построения визуальных моделей данных, для различных формулировок задач анали-

за. Показана возможность увеличения результативности визуального анализа в результате использования средств визуальной аналитики, обеспечивающих взаимодействие специалистов, имеющих различные области специализации и предварительной информированности.

Работа выполнена при поддержке Российского научного фонда, проект №18-11-00215.

## 8. Список литературы

1. Averchenkov V., Budylskii D., Podvesovskiy A., Averchenkov A., Rytov M., Yakimov A. Hierarchical deep learning: A promising technique for opinion monitoring and sentiment analysis in russian-language social networks // Proceedings of the First Conference, CIT&DS 2015: Creativity in intellectual technologies & data science. Volgograd, Russia, September 15 – 17, 2015. – Springer International Publishing, Switzerland, 2015, pp. 583–592.
2. Bondarev A.E., Galaktionov V. A. Multidimensional data analysis and visualization for time-dependent CFD problems // Program. Comput. Softw, Vol. 41, № 5, 2015, pp. 247–252.
3. Chen C. Top 10 unsolved information visualization problems // IEEE Comput. Graph. Appl, Vol. 25, № 4, 2005, pp.12–16.
4. Chen C., Czerwinski M. P. Empirical evaluation of information visualizations: An introduction // Int. J. Hum. Comput. Stud, Vol. 53, № 5, 2000, pp. 631–635.
5. Cook K. Mixed-initiative visual analytics using task-driven recommendations IEEE, 2015, pp. 9–16.
6. Pirolli P. Information Foraging Theory: Adaptive Interaction with Information // Inf. Foraging Theory, 2007, pp.1–28.
7. Pontis S., Blandford A. Understanding “influence”: An empirical test of the Data-Frame Theory of Sensemaking // J. Assoc. Inf. Sci. Technol, Vol. 67, № 4, 2016, pp.841–858.
8. Shklyar A. V., Zakharova A. A. Visual presentation of different types of data by



dynamic sign structures // *Sci. Vis.*, Vol. 8, № 4, 2016, pp.28–37.

9. Shklyar A., Zakharova, A., Zavyalov, D., Vekhter, E. *Visual detection of internal patterns in the empirical data* Springer Verlag, 2017., pp.215–230.

10. Барт Р. Риторика образа // *Избранные работы. Семиотика. Поэтика*. М., 1989, 616с.

11. Шкляр А. *Структурный подход к визуализации данных* Издательство МАИ-Принт (Москва), 2017, 609–611с.