



Рис. 1 Стандартное представление параметров образовательного процесса.

Проблемой практического анализа гетерогенных данных при проектировании и контроле результатов различных процессов является большой объем разнородных данных и ограниченность доступных ресурсов (вычислительных, временных, человеческих). Поиск новых инструментов анализа обусловлен необходимостью анализа большого числа связей между элементами общего процесса [2]. Традиционные средства обладают рядом недостатков, к которым следует отнести высокие требования к квалификации специалистов, участвующих в анализе, а также необходимость междисциплинарного взаимодействия между ними. Кроме того, улучшение достигаемых результатов подразумевает устранение внутренних противоречий в исследуемых данных, источником которых являются ошибки, допущенные при проектировании процессов. Субъективный характер таких противоречий затрудняет их обнаружение на основании предварительно сформулированных критериев.

## 2. Интерпретация и визуализация

Влияние на процесс интерпретации данных при решении прикладных задач оказывают объем данных и ряд дополнительных сведений, часть из которых не имеет прямого отношения к решаемой задаче. На этапе предварительного анализа исходных данных или в процессе решения поставленной задачи в распоряжении аналитика появляется множество дополнительных сведений (промежуточные решения, данные верификации, результаты классификации и кластеризации и т.д.). Кроме того, процесс исследования и интерпретации данных, требующий значительных временных ресурсов, сопряжен с необходимостью изучения новых — заменяющих или дополняющих прежние, данных [3].

Существенное влияние на достижение цели исследования могут оказывать: история получения данных, способы их сбора, наличие ошибок, частичное отсутствие данных и т.д. Метаданные, т.е. данные второго порядка, с точки зрения интерпретации, выполняют роль дополнительного источника сведений для исследователя и входят в состав гетерогенных данных. Таким образом, особенностью исследования гетерогенных данных, попадающих в распоряжение пользователя при решении задачи планирования, является необходимость сопоставления множества разнородных описаний отдельных элементов процесса, учет промежуточных версий решений, необходимость актуализации интерпретируемых данных в ходе их анализа. В качестве перспективного способа преодоления указанных трудностей в работе предложено применение средств визуализации данных, обладающих возможностью организации интерактивного взаимодействия между пользователем и исходными данными [4].

Интерпретация гетерогенных данных является процессом получения новой информации и характеризуется объемом необходимых для этого ресурсов. Это условие формирует требования к средствам интерпретации данных. Интерактивная визуализация, используемая для представления и интерпретации исходных данных, делает возможным использование потенциала визуального восприятия для компенсации указанных сложностей. Важной особенностью визуального восприятия является возможность сопоставления нескольких событий (вариантов), а также убедительность получаемых сведений, являющаяся основанием для принятия необходимых решений.

Визуальный образ данных может быть исследован различными способами и, следовательно, может дать пользователю ответы на разные вопросы. В простейшем случае, визуализация предоставляет подтверждение или опровержение уже сформулированному ответу на вопрос и не требует дополнительного осмысления результатов интерпретации. Это позволяет выделить иллюстративную роль визуализации, имеющую выраженную практическую направленность, например, демонстрация результатов решений задач анализа, полученных другими средствами. Для визуализации такого типа

характерны развитие в направлении автоматизации построения моделей, а также поиск форм визуального представления, сокращающих время интерпретации.

В работе представлены способы более сложного использования визуальных моделей, которые основываются на непосредственном взаимодействии исследователя и средства визуализации. Возможность изменения состояния визуальной модели в процессе ее использования соответствует построению управляемой модели данных и создает широкий круг ее преимуществ, к которым относится целенаправленное использование состояния модели для наблюдения за результатами управляющего воздействия [5].

### 3. Формальная постановка задачи

Анализируемый процесс является упорядоченным множеством элементов  $O = \{O_1..O_n\}$ , где  $O_i$  — элементарный процесс, определяемый набором уникальных свойств [6]. При этом, свойства элемента  $O_i$  делятся на два подмножества, на основании их функциональных различий:

$$O_i = \{P_i\} = \{In_x, Out_y\},$$

где подмножество  $\{In_x\}$  — требования к процессу, т.н. входящие связи элемента  $O_i$ , а подмножество  $\{Out_y\}$  — результаты процесса, т.н. исходящие связи элемента  $O_i$ . Определенные таким образом связи элемента  $O_i$  являются его основными характеристиками, обеспечивающими взаимодействие с другими элементами и достижение цели основного процесса.

В соответствии с данными определениями, цель основного процесса может быть представлена в виде композиции исходящих связей, характеризуемой некоторыми параметрами, например, количеством исходящих связей (результатов) и некоторыми весовыми коэффициентами. Таким образом, мощность множества  $O = \{O_1..O_n\}$  определяется предварительно определенной целью  $G = \{Out\}$  и доступными ресурсами.

Формальной целью решения сформулированной таким образом задачи планирования является поиск множества  $O$ , обеспечивающего заданные или наилучшие результаты при соблюдении всех граничных требований:

$$O = \{O_n : n = \min, G = \max\}.$$

В простейшем случае, когда свойства элементов  $O_i$  заданы, целью решения становится упорядочивание множества  $O$ , обеспечивающее выполнение граничных требований. В противоположном случае, решением задачи является новое множество  $O^* = \{In^*, Out^*\}$ .

### 4. Возможности пользователя

Возможности человеческого мышления разнообразны и недостаточно изучены [7], поэтому для визуализации данных выделены следующие:

- Адаптивность, т.е. готовность к включению новых сведений в систему уже существующих понятий и их категорий. Полезным следствием этого свойства является способность к восприятию быстро изменяющейся визуальной информации, в том числе, наблюдение за сценами, являющимися сложными наборами объектов и событий, а также выделение взаимодействий, в том числе новых.
- Избирательность внимания, которая исключает из набора перцептивных данных детали, не находящиеся в фокусе внимания. Это оптимизирует деятельность в условиях ограниченности ресурсов (например, временных).
- Двухступенчатый механизм восприятия, который, взаимодействуя с накопленным опытом, создает гибкий процесс оперативного осмысления визуального образа.

Необходимость использования возможностей восприятия, характеризуемых эмпирическими принципами, ставит перед разработчиком средств визуализации задачу определения набора характеристик пользователя, целенаправленно привлекаемых к

интерпретации визуального образа данных [8]. Выбор может быть осуществлен на основании предварительной оценки доступных ресурсов, к которым относятся:

- Характеристики потенциального пользователя, привлечение которых к процессу интерпретации данных не создает предпосылок для увеличения его продолжительности.
- Вычислительные ресурсы, позволяющие получать визуальные образы данных при условии соответствия требованию интерактивности взаимодействия с ними.
- Временные ресурсы, определяющие скорость построения и интерпретации визуальных образов данных.
- Дополнительные требования, следующие из постановки задачи исследования, в том числе: предварительная информированность пользователя, его квалификация, вероятные особенности восприятия и т.д.

Набор характеристик пользователя, привлекаемых к когнитивной интерпретации визуального образа исследуемых данных, может быть определен в результате обобщения существующих схем процесса визуализации [9]. На основании полученного набора выделены три группы действий пользователя средств визуализации, различающиеся по функциональному назначению и способам практической реализации:

- Наблюдение. Получение визуальной информации (восприятие цвета, пространства, движения, выделение групп, распознавание форм, знаков).
- Поиск. Выделение в исходной визуальной информации релевантных объектов и процессов (пространственное мышление, предварительная информированность, мотивация).
- Формулирование. Формулирование гипотезы ответа на вопрос исследования (опыт использования средств визуальной аналитики, способность к изучению и применению новых языковых систем), формализация новой информации.

Развитие технологий компьютерной визуализации и их непрерывное усложнение создает трудности при взаимодействии пользователя и разрабатываемых средств визуализации [10]. Это обстоятельство становится критически важным в ситуации, когда средства визуализации обеспечивают совместное участие в работе группы исследователей или являются способом обмена информацией между специалистами с различным уровнем подготовки или областью специализации. Возникает необходимость выбора между обучением пользователей использованию новых средств визуализации или привлечением уже существующих навыков визуальной коммуникации к интерпретации данных.

## **5. Средства визуализации для задачи планирования**

На основании сформулированных требований к средствам визуализации, необходимым для решения задачи планирования, разработана система интерактивного представления данных, входящих в описание произвольной образовательной программы. В качестве средства визуализации, предназначенного для решения поставленной задачи, предложена трехмерная визуальная модель, формирующая визуальный образ информационных объектов, входящих в исходные данные (

Рис. 1). Информационным объектом является элемент образовательной программы (учебная дисциплина, курс, раздел программы). Каждый такой объект является массивом данных (название, курс, продолжительность, объем, входящие требования, планируемые результаты), включающим переменные различного типа.

Разработано программное обеспечение, позволяющее получать интерпретируемый визуальный образ с привлечением визуализаторов пакета Autodesk 3ds Max. Алгоритм построения визуального образа реализован на языке Maxscript, поэтому обладает переносимостью и легко адаптируется к новым техническим возможностям визуализации. Интерфейс средства визуализации лишь частично использует среду Autodesk 3ds Max и может быть адаптирован к потребностям конкретного пользователя.

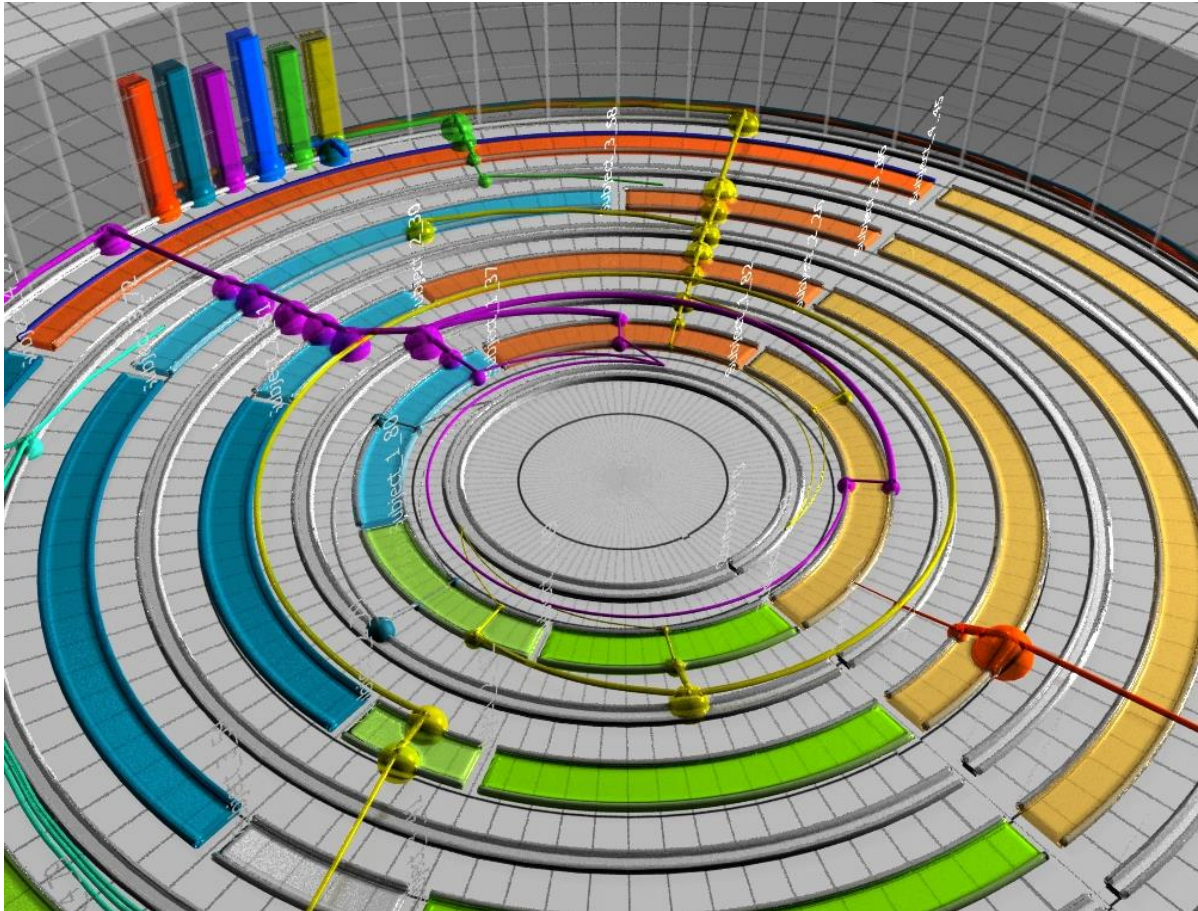


Рис. 2 Средство интерактивного представления данных, входящих в описание произвольной образовательной программы.

Обоснованная система интерактивного управления создает условия для постановки новых вопросов исследования и быстрого получения ответов, ускоряя достижение цели анализа. Следовательно, интерактивные особенности средства визуализации определяют последовательность и логику рассуждений исследователя.

Для сокращения периода обучения, необходимого для знакомства с новым инструментом интерпретации, предложено использование метафоры представления, опирающейся на традиционные приемы визуализации табличных данных (диаграммы, графики). В 3D пространстве средства визуализации задана цилиндрическая система координат, позволяющая сопоставить каждой точке пространства три величины: время обучения, нагрузку, результат. Масштабы единиц измерения по осям времени и результатов могут быть произвольными, нагрузка измеряется в процентах от максимально возможной. Возможность сравнения объектов реализована благодаря использованию цветового кодирования, правила которого могут быть изменены в соответствии с особенностями восприятия конкретного пользователя.

Радиальное направление оси времени позволяет визуализировать данные об образовательных программах любой продолжительности (бакалавриат, специалитет и магистратура). Этапы программы, соответствующие заданным временным интервалам (годы, семестры) разделены концентрическими круговыми элементами, используемыми для представления накопленных результатов. Каждый концентрический элемент является шкалой отсчета (0-100%) с общей для всех начальной точкой. Предложенная структура обеспечивает представление возрастающего числа результатов обучения без ухудшения общего восприятия данных.

Информационный объект предоставляет наблюдателю возможность интерпретировать визуальные атрибуты как значения соответствующих параметров: цвет является

атрибутом идентификации, размеры — нагрузки, положение — периоду обучения. Сведения о входящих требованиях и планируемых результатах обучения представлены в виде связей между информационными объектами. Атрибутами таких связей являются их направление и число, соответствующие исходным данным. В соответствии с особенностями предметной области, входящие (требования) и исходящие (результаты) связи имеют один тип, т.е. могут быть интерпретированы как создаваемые или развивающиеся компетенции.

Для визуального представления накапливаемых результатов обучения предложено одновременное использование двух выразительных средств: шкала оценки результативности и визуальное масштабирование связей. В первом случае становится возможным результативное использование трехмерного пространства визуальной модели, во втором — оценка результатов обучения по образовательной программе может происходить при интерпретации двумерной визуализации.

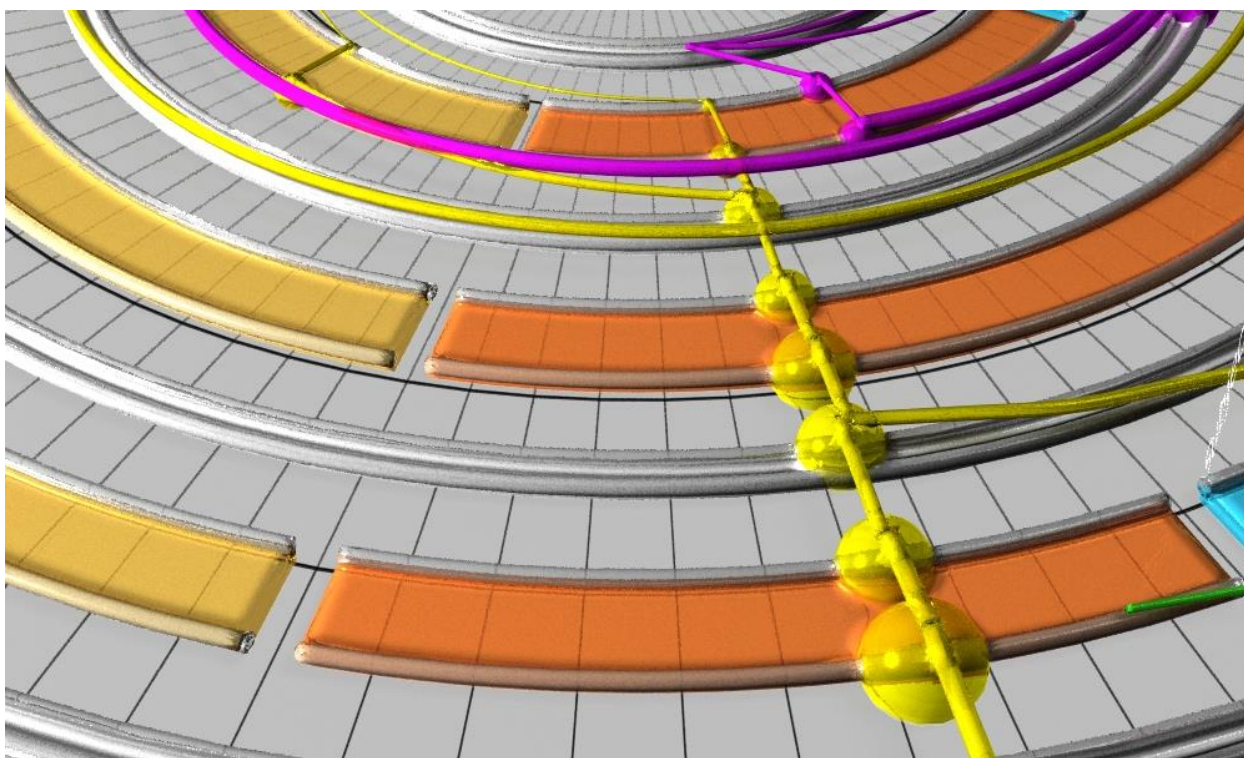


Рис. 3 Визуализация накапливаемых результатов

Разработан способ визуализации промежуточных и общих результатов в виде профилей суммирования результатов. Профиль имеет вид гистограммы, представляющей накопленные результаты в процентах от запланированных значений. Для упрощения образа исследуемых данных, элементы представления результатов могут быть временно исключены из визуальной модели.

Визуальные объекты, представляющие пользователю данные о достигаемых в исследуемом процессе результатах обучения, выполняют, кроме этого, одну из основных функций, состоящую в поиске вариантов характеристик отдельных элементов, удовлетворяющих цели планирования. Таким образом, предложено использование элементов визуального представления в качестве системы интерактивного управления состоянием средства визуализации. Создаваемый в этом случае интерфейс взаимодействия со средством визуализации становится основанием когнитивной интерпретации образов исходных данных с учетом индивидуальных особенностей мышления пользователя.

## 6. Противоречия в результатах планирования

Перспективными вариантами использования взаимодействия пользователя с разработанным средством визуализации можно рассматривать динамическую коррекцию планируемых результатов в результате обнаружения и устранения противоречий, допущенных на этапе планирования.

Наиболее быстро происходит поиск противоречий нескольких типов:

- Хронологическое несоответствие между входящими требованиями информационного объекта и достигнутыми результатами (Рис. 4). Противоречие возникает в случае, если на входе информационного объекта требуются данные, которые будут получены позже. Визуализируется как линия накапливаемых результатов, идущая в обратном направлении.

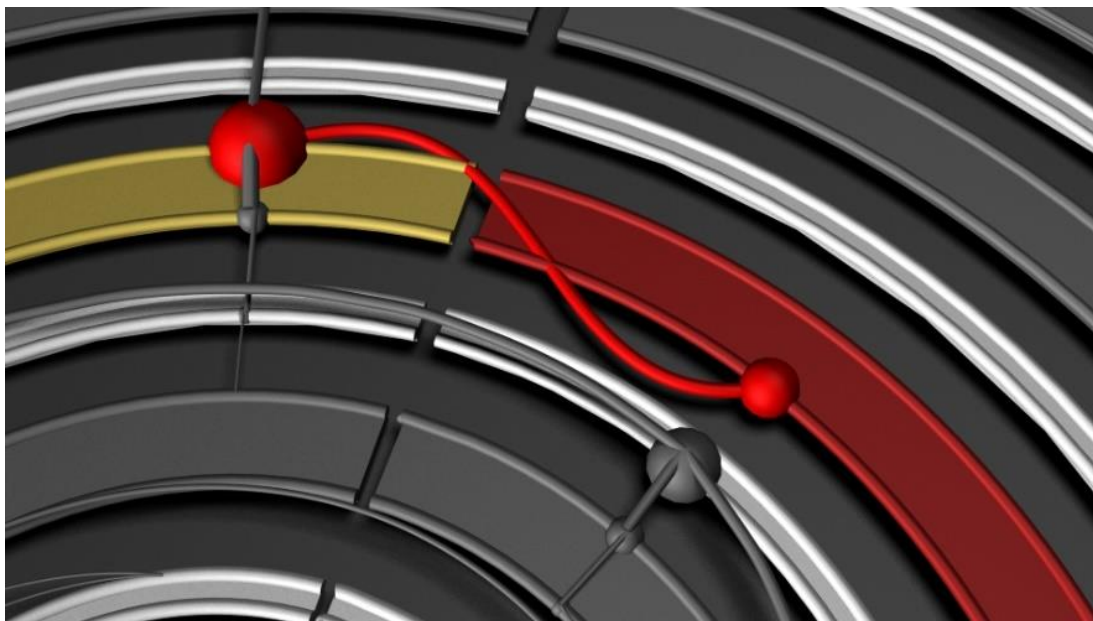


Рис. 4 Хронологическое несоответствие

- Отсутствие входных данных: входящие требования информационного объекта не обеспечены результатами, т.е. запрашиваемые данные отсутствуют. Обнаруживается как открытый вход информационного объекта (Рис. 5).

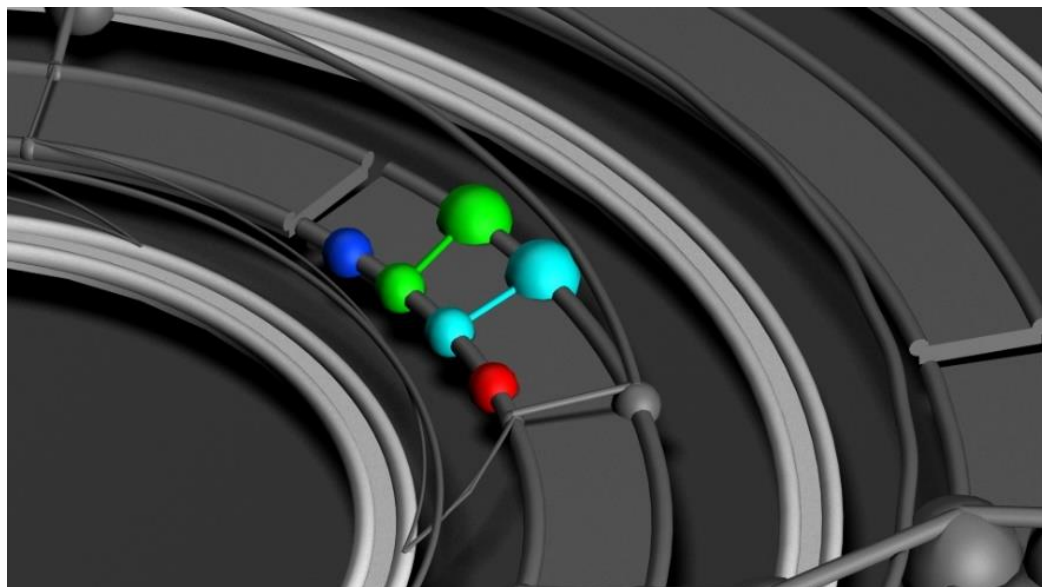


Рис. 5 Отсутствие входных данных

- Необоснованное использование существующих ресурсов. Результат интерпретации линий результатов, не представленных на выходе общего процесса.
- Деградация результатов. В предложенном средстве визуализации учитывается возможность снижения уровня достигнутого результата со временем. Противоречием считается наличие в итоговом профиле результатов, не учитывающих уровень их деградации.
- Дублирование результатов. Использование ресурсов для многократного получения схожих результатов. Визуальный образ подобного противоречия представляет собой расходящиеся линии накапливаемых результатов

#### Оценка результативности

Практическое использование разработанного средства визуализации позволило получить оценку преимуществ предложенного подхода по сравнению с традиционными способами интерпретации и верификации гетерогенных данных, содержащихся в документах, регламентирующих образовательные программы.

Table 1. Сравнение длительности этапов исследования данных.

<b>Параметр сравнения</b>	<b>Традиционный подход</b>	<b>Визуальный анализ</b>
<b>Обучение</b>	Предшествующий опыт	Менее 5 мин.
<b>Формирование общего представления</b>	15-20 мин.	Менее 2 мин.
<b>Поиск противоречий</b>	до 30 мин.	1-5 мин
<b>Изменение цели</b>	Невозможно	Интерактивно
<b>Управление данными</b>	Невозможно	10-20 мин.

Разрабатываемое средство визуальной аналитики образовательной среды может быть дополнено возможностью сохранения опыта использования, когда сохраняются уже сформированные варианты образовательных траекторий и соответствующие им реальные результаты. Расширены возможности опережающего планирования на различные по продолжительности периоды, основанные на анализе факторов, оказывающих влияние на отклонение реальных значений достижения профессиональных компетенций от запланированных в процессе проектирования образовательной программы.

## 7. Заключение

Использование средств визуализации в качестве инструмента оперативного исследования данных в задачах планирования является примером визуального анализа гетерогенных данных. В число получаемых преимуществ входят высокая скорость визуального восприятия, необходимая для одновременного сопоставления большого объема разрозненных фактов, а также возможность интерпретации данных не только в числовой форме.

Методика применения средства визуализации позволяет достигать цели исследования без интерпретации численных значений. Таким образом, предложен подход к решению задач указанного типа, опирающийся исключительно на средства визуализации. Средство визуализации имеет низкую зависимость от специфики предметной области задачи. Таким образом, опыт использования предложенного средства интерпретации может быть привлечен к решению других задач. Низкие требования к предварительной подготовленности пользователей обеспечивают возможность использования средства визуализации при отсутствии специальной подготовки при решении задач интерпретации данных, имеющих междисциплинарный характер.

Работа выполнена при поддержке Российского научного фонда, проект 18-11-00215.



## Список литературы

1. Guo H. [et al.]. A Case Study Using Visualization Interaction Logs and Insight Metrics to Understand How Analysts Arrive at Insights // IEEE Transactions on Visualization and Computer Graphics, 2016, Vol. 22, Issue 1, pp. 51–60. doi: 10.1109/TVCG.2015.2467613
2. Zakharova A.A., Vekhter E.V., Shklyar A.V., Zavyalov D.A. Visual Detection of Internal Patterns in the Empirical Data // // A. Kravets et al. (Eds.): CIT&DS 2017, Communications in Computer and Information Science, Vol. 754, pp. 215–230. doi: 10.1007/978-3-319-65551-2\_16
3. Manakov D., Averbukh V. Verification of Visualization // Scientific Visualization, 2016, Vol. 8, No. 1, pp. 58–94.
4. Blascheck T. [et al.]. VA2: A Visual Analytics Approach for Evaluating Visual Analytics Applications // IEEE Transactions on Visualization and Computer Graphics, 2016, Vol. 22, Issue 1, pp. 61–70. doi: 10.1109/TVCG.2015.2467871
5. Zakharova A.A., Vekhter E.V., Shklyar A.V. Methods of Solving Problems of Data Analysis Using Analytical Visual Models // Scientific Visualization, 2017, Vol. 9, No. 4, pp. 78–88. doi: 10.26583/sv.9.4.08
6. Zakharova A.A., Shklyar A.V., Rizen Y.S. Measurable Features of Visualization Tasks // Scientific Visualization, 2016, Vol. 8, No. 1, pp. 95–107.
7. Biederman I. Recognition-by-Components: A Theory of Human Image Understanding // Psychological Review, 1987, Vol. 94, No. 2, pp. 115–147.
8. Chen H. [et al.]. Uncertainty-Aware Multidimensional Ensemble Data Visualization and Exploration // IEEE Transactions on Visualization and Computer Graphics, 2015, Vol. 21, Issue 9, pp. 1072–1086. doi: 10.1109/TVCG.2015.2410278
9. Van Wijk J.J. The Value of Visualization // Proceedings of VIS 05. IEEE Visualization, 2005, doi: 10.1109/VISUAL.2005.1532781
10. Chen C. Top 10 Unsolved Information Visualization Problems // IEEE Computer Graphics and Applications, 2005, Vol. 25 Issue 4, pp. 12–16. doi: 10.1109/MCG.2005.91

# Using Visualization Tools to Search for Contradictions in Planning Results

A.A. Zakharova<sup>1,A</sup>, E.V. Vekhter<sup>2,B</sup>, A.V. Shklyar<sup>3,B</sup>

<sup>A</sup> Bryansk State Technical University

<sup>B</sup> Tomsk Polytechnic University

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

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

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

## **Abstract**

When designing most types of activities, as well as when needed to audit the results achieved, a problem arises to analyze simultaneously data included in the descriptions of individual elements of this activity. The purpose of such analysis may be to identify contradictions among disparate information descriptions of processes included in the general activity, as well as to search for possible ways of improving the results achieved.

The paper presents the results of a study of visualization potential in the analysis of large volumes of heterogeneous data. The results obtained have been used in the development of visual research tools which help to improve the design of complex processes and to solve the applied problem of planning the educational process.

**Keywords:** visual analytics, visual model, data analysis, visual interpretation, visual perception.

## **8. References**

1. Guo H. [et al.]. A Case Study Using Visualization Interaction Logs and Insight Metrics to Understand How Analysts Arrive at Insights // IEEE Transactions on Visualization and Computer Graphics, 2016, Vol. 22, Issue 1, pp. 51–60. doi: 10.1109/TVCG.2015.2467613
2. Zakharova A.A., Vekhter E.V., Shklyar A.V., Zavyalov D.A. Visual Detection of Internal Patterns in the Empirical Data // // A. Kravets et al. (Eds.): CIT&DS 2017, Communications in Computer and Information Science, Vol. 754, pp. 215-230. doi: 10.1007/978-3-319-65551-2\_16
3. Manakov D., Averbukh V. Verification of Visualization // Scientific Visualization, 2016, Vol. 8, No. 1, pp. 58–94.
4. Blascheck T. [et al.]. VA2: A Visual Analytics Approach for Evaluating Visual Analytics Applications // IEEE Transactions on Visualization and Computer Graphics, 2016, Vol. 22, Issue 1, pp. 61–70. doi: 10.1109/TVCG.2015.2467871
5. Zakharova A.A., Vekhter E.V., Shklyar A.V. Methods of Solving Problems of Data Analysis Using Analytical Visual Models // Scientific Visualization, 2017, Vol. 9, No. 4, pp. 78–88. doi: 10.26583/sv.9.4.08
6. Zakharova A.A., Shklyar A.V., Rizen Y.S. Measurable Features of Visualization Tasks // Scientific Visualization, 2016, Vol. 8, No. 1, pp. 95–107.
7. Biederman I. Recognition-by-Components: A Theory of Human Image Understanding // Psychological Review, 1987, Vol. 94, No. 2, pp. 115–147.
8. Chen H. [et al.]. Uncertainty-Aware Multidimensional Ensemble Data Visualization and Exploration // IEEE Transactions on Visualization and Computer Graphics, 2015, Vol. 21, Issue 9, pp. 1072–1086. doi: 10.1109/TVCG.2015.2410278

9. Van Wijk J.J. The Value of Visualization // Proceedings of VIS 05. IEEE Visualization, 2005, doi: 10.1109/VISUAL.2005.1532781
10. Chen C. Top 10 Unsolved Information Visualization Problems // IEEE Computer Graphics and Applications, 2005, Vol. 25 Issue 4, pp. 12–16. doi: 10.1109/MCG.2005.91