

ИСКУССТВОВЕДЕНИЕ

УДК 004.91

DOI: 10.17223/22220836/21/9

Н.С. Бажанов

О СИСТЕМЕ ПОИСКА ИНФОРМАЦИИ В ТЕКСТАХ ИСТОРИЧЕСКОГО МУЗЫКОЗНАНИЯ

В статье представлена система поиска информации, состоящая из информационной потребности, поискового запроса, поисковой машины, информационного объекта, результата поиска. В составе информационной потребности раскрывается информационный парадокс, когда для полноты поиска требуется законченность представлений о научной проблеме, что на начальной стадии исследования невозможно. В качестве выхода из парадокса предлагается возможный инвариант строения научного текста в историческом музыкознании. Цель статьи – совершенствование технологии информационного поиска при работе с научными текстами.

Ключевые слова: информационный поиск, научная информация, анализ текста, система поиска информации, компьютерное ассистирование в редактировании текстов, интеллектуальный анализ текста, музыкознание.

Информационный поиск¹ – важная часть компьютерных технологий и вместе с тем междисциплинарная область науки, включающая когнитивную психологию, информатику, лингвистику, семиотику, логику и некоторые другие дисциплины² [1].

Необходимость совершенствования технологии поиска информации вызвана рядом причин. Прежде всего, это информационные характеристики современной культуры и лавинообразный рост информации³.

¹ Термин «информационный поиск» (англ. «information retrieval») ввёл американский математик К. Муэрс. Он заметил, что побудительной причиной такого поиска является информационная потребность, выраженная в форме информационного запроса. К объектам информационного поиска К. Муэрс отнес документы, сведения об их наличии и (или) местонахождении, фактографическую информацию. Под информационной поисковой системой ИПС (англ. «information retrieval system», IRS) понимают системы, обеспечивающие поиск информации, хотя их конфигурации и функции могут быть весьма различны.

² Существует комплексное научное направление исследований в лингвистике – обработка естественного языка компьютерными технологиями. Частью этой проблемы является автоматизированная обработка текста, в свою очередь состоящая из информационного поиска, машинного перевода, автоматизированного реферирования, вопросно-ответных систем, ведения диалога (с виртуальными собеседниками, помощниками), интеллектуального анализа текста (text mining).

³ В научной литературе существует более 400 определений термина «информация». Очень часто содержание этого термина лишь частично оговаривается, отчасти конвенционно, отчасти объясняется контекстом употребления в научном тексте.

Информация – это то, что не существует без субъекта, который её воспринимает.

Сообщение несёт информацию для человека, если содержащиеся в нем сведения являются новыми и понятными.

Информация – это «мера структурного разнообразия» (У. Эшби).

Информация это информация, не вещество и не энергия, и этого достаточно! (Н. Винер).

Сегодня уже измеряют не объем информации, а скорость его удвоения. Так, в первый раз для удвоения информации в культуре человечеству понадобилось 1750 лет (точка отсчета – Рождество Христово). Следующее удвоение наступило в 1900 г., а затем для удвоения всей информации на планете понадобилось уже 50 лет (1950 г.). В 2000 г. цикл удвоения объема информации составлял всего лишь 5 лет. Во второе десятилетие XXI в. считается, что информация удваивается примерно за 180 дней [2].

Всемирный книжный фонд удваивается каждые 10–15 лет, число телефонных каналов – каждые 11 лет, число автоматизированных баз данных увеличивается в 10 раз за 10 лет. В научных работах растут ссылки на литературу. В гуманитарных дисциплинах, в современных статьях список литературы насчитывает 40–60 наименований на 1 печатный лист текста – настолько велик научный контекст исследований.

Рост объема информации – реалии современной жизни. По данным IDC (InternationalDataCorporation – аналитическая фирма, исследующая рынок информационных технологий), на сегодняшний день проанализировано менее 1% всей имеющейся информации.

Хорошо известная всем «озоновая дыра» над Антарктидой была обнаружена американским метеорологическим спутником еще в 1979 г., но информация утонула в архиве, насчитывающем около 3 млн видеолент. И только спустя 7 лет английские ученые расшифровали видеоматериалы, им и досталась честь открытия.

Итак, наряду с накоплением информации действуют и нарастают отрицательные факторы для её поиска, восприятия и понимания. Возникает диффузия информации. Она размывается по языкам, по сайтам, по отраслям деятельности и дисциплинам науки. С ростом информации объективно затрудняется поиск необходимых сведений.

Процесс дробления и специализации научных дисциплин, дифференциация знаний приводят к сужению тематических границ профессиональной информационной потребности. Все более узким специалистам требуется все более специализированная информация. Необходимо в «стоге сена найти не иголку, а отдельную молекулу». Отношения между найденным и отсеянным приближается к дробным величинам, близким к нулю, что объективно затрудняет процесс поиска.

В качестве выхода из «кризиса поиска информации» уже предлагаются двухэтапные процедуры, когда «первый этап – это предварительный поиск и отбор информации в тематические базы данных, а второй этап – это поиск нужной информации... в сетевых или локальных полнотекстовых базах» [3].

В гуманитарных областях знания практически отсутствует деятельность, концентрирующая информацию по её спецификации, – дайджесты, библио-

Информация – сведения, воспринимаемые человеком и (или) специальными устройствами как отражение фактов материального или духовного мира в процессе коммуникации (ГОСТ СИБИД 7.0-99. «Информационно-библиотечная деятельность. Библиография. Термины и определения»).

Информация – сведения (сообщения, данные) независимо от формы их представления («Об информации, информационных технологиях и о защите информации». Статья 2. Федеральный закон РФ от 27 июля 2006 г. № 149-ФЗ. Цит. по [1]).

графические обзоры, сайты с кратким библиографическим описанием научных изданий. Это положение особенно касается научных дисциплин искусствоведения.

Поиск информации – важная часть любой научной работы. Не будем касаться случаев, когда исследование предполагает экспериментальное извлечение из небытия новой информации. Это научно-исследовательский поиск сведений, когда неизвестная, введенная в обиход научной дисциплины информация получена в результате кропотливой экспериментальной, экспедиционной или архивной работы ученого. С точки зрения новизны это первичная информация.

В настоящей статье речь идет о поиске информации, уже *существующей* в человеко-машинных информационно-поисковых системах, о вторичном поиске сведений. Рассматривается поиск информации, ограниченный полнотекстовыми базами.

Чтобы совершенствовать поиск информации, необходимо более полно и целостно представить структуру информационного поиска.

Успешность поиска научной информации (результат) зависит от многих факторов:

- от уровня осознания информационной потребности;
- технических характеристик поисковой машины;
- адекватности и изощренности запроса на поиск информации;
- свойств информационного объекта (базы данных), в котором осуществляется поиск.

Тогда в нашем варианте система информационного поиска будет состоять из пяти элементов: *информационная потребность*, *поисковая машина* и её возможности, *запрос на поиск информации*, *информационный объект* или массив, в котором осуществляется поиск, и сам *результат* поиска.

Кратко характеризуем эти элементы и их функции в системе поиска.

Информационная потребность – элемент системы поиска информации, в котором исследователь участвует в наибольшей степени. Потребность, необходимость информации формируется в сознании исследователя. Именно в этой части системы сказывается его искусство вопрошания. («Всякое знание проходит через вопрос», – утверждал Ганс Георг Гадамер в своем труде «Об искусстве вопрошания») [4].

Мера осознания информационной потребности может быть разная. Представим информационную потребность в виде иерархических уровней, когда каждый последующий включает и дополняет предыдущие.

1. В таком случае первый и самый элементарный уровень осознания информационной потребности представлен *необходимостью в научной литературе* по теме исследования. Как правило, это библиографические описания книг, статей, ссылки на сайты, в названии которых будет представлена тема исследования. Список найденной литературы должен быть наиболее полным и новым по данной проблеме. В соответствии с современным состоянием изучения научной проблемы выстраиваются идея, концепция исследования, формируется исходная аксиоматика, терминологический аппарат. Но даже на этой тривиальной ступени должна быть задействована научная содержательная аналитика в части поиска литературы *по синонимам* запроса или по род-

ственным понятиям. Тематику поиска необходимо выразить словами, содержание смыслов облечь в вербальную форму. Серьезную аналитическую проблему представляет *формулирование запроса* в компьютерной системе поиска информации, чтобы результат был *наиболее полным и точным* по содержанию.

2. Потребность в информации может быть выражена в виде необходимости сведений о *терминологическом аппарате* исследования. Тогда формируется информация о комплексе родственных понятий и терминов, исчерпывающе описывающих основную проблему исследования. К терминам и понятиям примыкает функциональная система отношений, взаимодействий между ними. Запросы и поиск должны находить вербальные эквиваленты таких отношений между терминами исследования или между терминами исследования и категориями (основными фундаментальными терминами) научной дисциплины.

3. Следующий уровень информационной потребности в системе поиска – необходимость найти смысловые связи в большом корпусе текстов, логические взаимодействия между *суждениями основного научного текста и контекста информационного окружения*¹. Может существовать потребность в *умозаключениях и выводах*, извлеченных системой поиска из корпуса тематически однородных текстов. Итак, этот уровень следует обозначить как *логический*, состоящий из суждений, умозакключений, обобщений, выводов.

4. Кроме того, различают потребность в фактографической и концептуальной информации. В первом случае это сведения о фактах: датах, именах, литературных источниках, т. е. справочная информация. Вместе с тем иногда необходимы оценки, интерпретации событий, заключения на высоком уровне абстрагирования, изложения теорий или концепций. В таком варианте система поиска должна выявить концептуальную информацию, что наиболее типично для современных научных исследований.

5. Наконец, большие научные полнотекстовые базы концентрированной тематики позволяют вести так называемый глубинный анализ для выявления *существующих, но необнаруженных* взаимосвязей, тенденций между логико-теоретическими элементами научного текста.

Следовательно, во-первых, потребность в информации уже разделяется на функционально различные части. Очевидны потребность в литературе по теме исследования, потребность в фактах, потребность в понятиях, терминах, категориях, суждениях, умозакключениях, взаимосвязях и выводах. Вместе с тем существует потребность в поиске смысловых элементов самого научного текста. В этом случае поиск информации становится частью построения самого научного текста, частью совершенствования его формы и содержания.

Во-вторых, становится понятным, что поиск информации не разовая конечная функция, а непрерывная, постоянная технология компьютерной работы с научными текстами.

Поисковая машина. Возможности поисковой машины должны обеспечить достаточно быстрое завершение обработки информационного объекта. Кроме

¹ Массив научной информации в границах любой тематики сегодня настолько огромный, что в ходе изложения научного «сюжета» исследования практически по каждому суждению, умозакключению может быть найдена информация в полнотекстовых базах.

того, машина (программа поиска) должна предоставить возможность исследователю весьма *гибко и изощренно сформировать запрос*, адекватный информационной потребности. Перед поисковыми машинами ставится ряд лингвистических задач – морфологический анализ, разрешение лексической многозначности и т. д. Синтаксис запроса и его разнообразие определяются возможностями поисковой программы, так же как и текстовые форматы (txt., doc., pdf), с которыми может работать поисковая машина. К сожалению, в гуманитарных исследованиях вынужденно используются бытовые поисковые машины широкого пользования, а не специализированные поисковики, ориентированные на содержание научной дисциплины.

Запрос на поиск информации. Все технологические особенности, способы, методы поиска в итоге приводят к формированию *текста запроса* для поиска информации. Запрос обладает свойством финального компонента в рядоположенной цепочке элементов системы поиска информации. Это может быть не только ключевое слово, но и группа синонимов, перекрывающих семантическое поле, или сложная, составная логико-вербальная конструкция. Диапазон синтаксиса запроса простирается от отдельного слова до множества слов совместно со знаками булевой логики, вплоть до поиска фраз из нескольких слов с указанием пробелов, разрывов между ними и заданных последовательностей.

Сводя разнообразие содержания информационной потребности к лаконичной форме запроса, исследователь должен проявить интеллект, научное знание и опыт, а также отчасти быть лингвистом. Функциональная задача запроса в системе поиска информации – короткой вербальной конструкцией «отсеять» достаточно объемное содержание информационной потребности. Иногда в один «заход» поиска сделать это не удастся, и процедура поиска становится многоступенчатой. Помочь исследователю могут справочные материалы: различные электронные тезаурусы, словари и прежде всего – словарь синонимов.

Различают *язык* поискового запроса (русский, английский и т. д.), *синтаксис запроса* (ключевые слова и логические знаки), *объект запроса* (текстовой документ, изображение, видео, музыка, звучание и т. д.).

Запросы бывают: единичные и множественные; однократные и многократные; суммативные и системные; монологичные и диалогичные.

Единичные – состоят из одного слова;

Множественные – несколько слов (включая булеву логику между ними), синонимические поля; семантические поля;

однократные – один запрос = один ответ;

многократные – много запросов = множество ответов, но в одной семантической зоне;

суммативные – много запросов несистемного, разрозненного характера;

системные – множество запросов, ответы на которые взаимосвязаны, представляют собой целостность, систему понятий;

монологичные – запросы, составленные вне зависимости от полученных ответов;

диалогичные – когда каждый последующий запрос учитывает предыдущий ответ;

односмысловые – запросы для поиска единичного смысла;

взаимодействующие – запросы на поиск взаимодействия (отношения) между смыслами.

Возможны также смешанные виды запросов.

По степени совпадения запроса и ответа различают запросы на буквально точное совпадение и запросы на произвольное соответствие, когда один искомый смысл выражен группой синонимов.

Информационный объект. («Очень трудно найти в темной комнате черную кошку, особенно, если её там нет». Конфуций.) Понятно, что искомая информация должна потенциально присутствовать в том массиве, в котором организуется её поиск, и этот информационный объект должен обладать достаточно большим объёмом. Чем больше массив, тем преимущества компьютерного поиска информации реализуются наиболее полно. Вот почему столь велика потребность в *полнотекстовых* предметно-ориентированных базах данных.

Необходимо различать несколько типов объектов для поиска, поскольку типология информационного массива, безусловно, влияет на текст запроса. Первый тип образует *смешанная*, тематически не сепарированная *информация*. Ко второму типу будем относить подготовленные информационные массивы. В них присутствует хотя бы общее тематическое единство. По форме объекты для поиска соответствуют типу файлов – тексты в различных форматах, изображения, музыка, видео.

Результат поиска необходимо «унести с собой». Поисковая машина должна позволять удобно копировать результаты поиска для дальнейшего использования в текстах научной деятельности.

Влияние на состояние системы оказывают не только характеристики её элементов (информационная потребность, запрос, поисковая машина, информационный объект, результат), но и отношения между этими элементами. Для поиска значимы взаимодействия всех элементов структуры между собой, хотя бы на уровне бинарных отношений, т. е. парных взаимодействий элементов. Например, от качества информационной потребности зависит синтаксис запроса, но одновременно текст запроса ограничивает информационную потребность. Вербализация запроса в виде ключевой конструкции уточняет информационную потребность. И так во всех парных отношениях происходит взаимовлияние элементов.

Парадокс полноты поиска информации. Поиск информации возникает не на пустом месте, а в некотором недостаточном, предварительном информационном контексте, что можно обозначить как «информационное пространство проблемы». Оказывается, что на начальном этапе исследования наиболее полно представить потребность в информации возможно, лишь хорошо представляя *структуру проблемы* в целом¹. Поиск литературы, выписки цитат, ограничения объекта изучения, развитие концепции, формирование аксиоматики происходят в начале исследования. Полное осознание и видение строения научного текста возникает *на завершающей стадии* работы, а наибольшая потребность в информации – в её начале. Таким образом, в рамках

¹ Мы искали информацию спонтанно (отчасти в случайном порядке, лишь ограничиваясь темой исследования), потом разложили информацию по полочкам. Почему же мы не искали информацию исходя из названий тех самых полочек, исходя из структуры научного исследования?

системы поиска формировать на высоком уровне полноты информационную потребность возможно лишь на завершающей стадии научного исследования. Парадокс поиска информации проявляется в информационной потребности и неизбежно влияет на все остальные элементы системы поиска информации. Если методология – *это способ представить исследование уже законченным*, то выход из данного парадокса – методологический.

Следовательно, для успешного поиска информации возникает необходимость в представлении некоего содержательного¹ обобщенного структурного инварианта многих научных работ в данной отрасли знания.

Как выглядит инвариант научных исследований в историческом музыкознании? В большинстве научных работ фигурируют следующие повторяющиеся элементы.

Наряду с *именами* композиторов, музыкантов обсуждаются их *атрибуты*: характеры, психологические портреты, биографические сведения. В большинстве исследований содержатся атрибуты пространства и времени: путешествия персонажей и места проживания, сопоставления хронологические и исторические. *Имена в контексте* пространства и времени – такое название можно дать этой части.

Обязательной частью любой работы будут музыкальные *произведения*. Их свойства, форма, содержание, история создания, исполнения, стиль, жанр, выразительные средства представлены в этой части. В общем смысле здесь присутствуют произведение, *атрибуты* произведения и его *контексты*.

В качестве примера в таблице приводится структура проблемы в области исторического музыкознания. В первом столбце показаны элементы научного текста – категории 1-го уровня, во втором столбце содержатся подкатегории, те понятия, на которые дробятся категории 1-го уровня, в третьем столбце – элементы, из которых состоит 2-й уровень. Получается структура с тремя иерархическими уровнями вложения.

Т а б л и ц а

| Категории 1-го уровня | Категории 2-го уровня | Категории 3-го уровня | Запрос на поиск |
|--------------------------|--------------------------|---|--|
| 1 | 2 | 3 | 4 |
| Персоналии | | Имена (нахождение заглавной буквы внутри предложения) | (не точка, пробел, заглавная от А до Я) «[!] [А–Я]» или в латинице «[!] [А–Z]» |
| Биографические сведения | Характеристики | Иконография | «N» выглядел, портрет, художник, изображен, внешность, наружность, лицо, лик, облик, фигура... |
| | | Психологический портрет | «N» был, характер, натура... |

Окончание таблицы

¹ В литературе методологического плана существует большое множество работ, раскрывающих формальную общую структуру исследования: постановка цели и задачи, научная проблема, материал, методология и т.д. Однако общая логическая структура научного текста рассматривается крайне редко.

| 1 | 2 | 3 | 4 |
|---------------------------|------------------------------------|--|--|
| Биографические сведения | Места проживания | Города, страны | «N» жил в... проживал, снимал квартиру, переехал, переселился... |
| | События | Жизни | родился, учился, женился... |
| | | Встречи | Встретились, познакомились... |
| | | Путешествия | Переехал, переселился, отправился, перебрался, путешествие... |
| | | Истории, свидетельства | – свидетель, произошло, случилось, N рассказывает... |
| | Учитель – ученик, сочинение музыки | Обучение, методика | «N» учился, брал уроки, обучал, преподавал... |
| Произведения | Содержание | Контекст сочинения | Название произведения |
| | | История сочинения | Название произведения |
| | Форма | Особенности выразительных средств, приемов | – соотношение частей, строение произведения, форма, содержание, прием, динамика, ритм, метр, агогика, динамический, ритмический, метрический, агогический... |
| | Стиль | Особенности выразительных средств, приемов | – атрибуты стиля, стилевой, характерный, стилистический, динамика, ритм, метр, агогика, динамический, ритмический, метрический, агогический... |
| | Жанр | Особенности выразительных средств, приемов | – жанровый, соната, симфония, квартет,... |
| Логические единицы текста | Имена | | ? |
| | Суждения | | ? |
| | Определения | | Называется, это, будем считать, представляет собой... |
| | Выводы | | Итак, следовательно, суммируя, ...вывод, заключение... |

Любой информационный поиск исходит из двух допущений. Считается, что:

а) документы, необходимые пользователю, содержат некоторые объединяющие признаки; б) пользователь способен указать этот признак. Относительно простых случаев поиска информации эти условия можно реализовать, но для сложного структурно-логического полнотекстового поиска выполнить эти условия удастся не всегда. Там, где оформить запрос на поиск затруднительно, стоят знаки «?».

Приведем пример поиска имен собственных в текстовом массиве. В таблице (строка 2, ячейка 4) показан запрос для редактора Word на поиск собственных имен. Этот сервис включается в «Найти/расширенный поиск / подстановочные знаки / найти в (основной документ)». Автоматически во всем тексте выделяется внутри предложения сегмент «не точка, пробел, заглавная буква». Однако при таком запросе не удастся автоматически отделить в поиске слова «Моцарт» и «Зальцбург» внутри предложения. Смыслы этих слов читатель различает по контексту предложения. В данной конфигурации запроса слово «Моцарт» в начале предложения будет пропущено. Следовательно, наиболее трудны для поиска случаи, когда смыслы, объединяющие

тексты или их фрагменты, извлекаются не из значения слов, а из контекста их употребления. Заполнить пустые ячейки 4-го столбца таблицы возможно на основе лексического анализа соответствующих фрагментов текста заданного содержания. Но это уже не концептуальная, а технологическая задача.

Доступ к фрагментам текста с заданным содержанием позволит организовать сканирование и отбор пользователем смыслов из огромных текстовых массивов. Такая технология может вывести работу исследователя на иной интеллектуальный и аналитический уровень.

Существует информационный поиск и информационный анализ текста. В отличие от машинного перевода, озвучивания текста и других направлений автоматической обработки текстов (АОТ) в данном случае в центре обработки текста стоит человеко-машинная система компьютерного ассистирования. Ни о какой автоматизированной, не зависящей от человека системе речь не идет. Человек не исключается из действия системы, а, наоборот, активно *включается*, участвует в ней, образуя важнейшую, ведущую часть, взаимодействующую с возможностями компьютерных технологий.

На какой-то стадии освоения информационных технологий работы с текстом поиск информации переходит на иной уровень. Чисто внешне запрос, поиск, результат поиска остаются как основные части технологии, но меняется их функция. Как средство работы с информацией они те же, но цели становятся иными. На смену поиску информации приходят поиск информационного *сопоставления*, обнаружение *информационных взаимосвязей*, выявление *структуры (строения)*, *информационное обеспечение для обобщения* и т.д.

Реализация этих функций в системе поиска обеспечивает высокий уровень взаимодействия исследователя и компьютерной технологии поиска информации. Такое взаимодействие принято называть ассистированием (computerassisted) [5].

Выводы. Поиск информации не разовая конечная функция, а непрерывная, постоянная технология цифровой компьютерной работы с научными текстами. В информационных поисковых системах (ИПС) стоит важная задача постепенного перехода от поиска информации к информационной аналитической работе с научным текстом.

Поиск информации представляет собой систему (состоящую из информационной потребности, запроса на поиск информации, поисковой машины, информационного объекта, результата поиска), между элементами которой существуют множественные *взаимосвязи*. Успешность поиска зависит от состояния всей системы в целом.

В настоящий момент активно формируется современная технология обработки и работы с научными текстами. Такая технология базируется на инварианте и общей логико-содержательной структуре множества научных текстов в отдельной научной дисциплине.

Поиск информации, обработка и работа с научными текстами наиболее эффективны в гуманитарных дисциплинах, где текстовой информационный объект наибольший по количеству информации, где отсутствует обобщенный формализованный язык (формулы), а смыслы выражены вербально.

В более обобщенном плане необходимо создавать некую новую компьютерную, *человеко-машинную технологию работы с научным текстом*. Вектор её развития – от поиска информации к современной технологии информационной работы с научными текстами.

Литература

1. Бекман И.Н. Информатика [Электронный ресурс] : курс лекций. Лекция 2. Информация. URL: <http://profbeckman.narod.ru/InformLekc.files/Inf02.pdf> (дата обращения: 08.01.2015).
2. Тоффлер Э. Революционное богатство: как оно будет создано и как оно изменит нашу жизнь / Э. Тоффлер, Х. Тоффлер. М. : АСТ, 2008. 570 с.
3. Захарченко В. Программы поиска информации в полнотекстовых базах данных [Электронный ресурс]. URL: http://www.mbdsoft.ru/articles_files/sereview.htm (дата обращения: 08.01.2015).
4. Гадамер Г.Г. Об искусстве вопрошания [Электронный ресурс] // Гадамер Г.Г. Истина и метод. Основы философской герменевтики. М., 1988. (Из классического наследия западной философии XX века). URL: <http://bookre.org/reader?file=332135> (дата обращения: 08.01.2015).
5. Center for Computer Assisted Research in the Humanities at Stanford University [Электронный ресурс]. URL: <http://www.ccarh.org> (дата обращения: 08.01.2015).

Bazhanov Nikolaj S. Novosibirsk State Conservatoire (Novosibirsk, Russian Federation).

E-mail: Bazhanov_Nikolaj@mail.ru

DOI: 10.17223/22220836/21/9

ABOUT THE RETRIEVAL SYSTEM IN THE TEXTS OF HISTORICAL MUSICOLOGY

Key words: *information retrieval, scientific information, text analysis, search engine for information, musicology.*

The article presents information search system, consisting of information needs, search, search engine, information of the object of the search result. Between the elements of the system, there are multiple relationships. The success of a search depends on the State of the system as a whole.

Composed of information needs revealed the information paradox, when to complete the search requires completeness notions about scientific problem that at the initial stage of research is impossible. A common awareness of the structure of scientific work should be preceded by a search of the literature and quoted material research, but this realization occurs only at the final stage. As you exit paradox proposed possible invariant structure of scientific text in historical musicology.

In the texts of historical musicology are present in the first-level Category: Personalities, names; Biographical information; Works; Logical units of text (Names, judgment, determination, conclusions). Second-level categories: characteristics; Place of residence; Events; Teacher-student, Especially composing music; Content works (form, style, Genre). Third level categories: iconography; Psychological portrait; City, country of residence; Meeting; Travel; History of the certificate; Learning methodology; Composing contexts; History essays; Especially expressive means, techniques. In article the data categories are represented in the form of a table where one column includes search keywords listed categories.

Information retrieval, processing and working with scientific texts are most effective in humanitarian disciplines, where textual information object largest quantity information where there is a generalized formal language (the formula), and meanings are expressed verbally. Most humanitarian texts expressed words meanings contained in the uniform text, without the intervention of a specialized formula language.

At some stages of development of information technologies of work with the text search information moves into a different phase. Superficially, the search query, the search results remain as major pieces of technology, but changing their function. As a means of information, they are the same, but the goals become different. On the change of information search, comes its mapping, information discovery, relationships, identification of structure (structure) information for synthesis, etc.

In a more generic sense, you need to create some new computer, man-machine technology with scientific text. Vector of development-from information retrieval to modern information technology with ongoing scientific texts.

References

1. Bekman, I.N. (n.d.) *Informatika* [Computer science]. [Online] Available from: <http://profbeckman.narod.ru/InformLec.files/Inf02.pdf>. (Accessed: 8th January 2015).
2. Toffler, A. & Toffler, H. (2008) *Revolutsionnoe bogatstvo: kak ono budet sozdano i kak ono izmenit nashu zhizn'* [Revolutionary Wealth: how it is created and how it will change our lives]. Translated from English by M. Sultanova & N. Tsykun. Moscow: AST.
3. Zakharchenko, V. (n.d.) *Programmy poiska informatsii v polnotekstovyykh bazakh dannykh* [Information retrieval programs in full-text databases]. [Online] Available from: http://www.mbdsoft.ru/articles_files/sereview.htm. (Accessed: 8th January 2015).
4. Gadamer, G.G. (1988) *Istina i metod. Osnovy filosofskoy germenевtiki* [Truth and Method. Fundamentals of philosophical hermeneutics]. Translated from German. Moscow: Progress.
5. Stanford Humanities Centre. (n.d.) *Center for Computer Assisted Research in the Humanities at Stanford University*. [Online] Available from: <http://www.ccarh.org>. (Accessed: 8th January 2015).