

Исходя из проведенного анализа, можно утверждать, что метод внедрения информационного комплекса будет эффективнее.

### **Заключение**

В ходе работы была поставлена задача, решение которой поможет ускорить процесс разработки и позволит организовать рабочий процесс в студии создания сайтов. После проведения сравнительного анализа было принято, что лучше всего данную задачу можно решить с помощью внедрения в компанию информационного комплекса. Информационный комплекс был выбран по необходимым критериям среди предложенных на рынке. При проведении оценки работы нескольких схожих по своим возможностям сервисов, было решено в качестве информационного комплекса выбрать приложение JIRA.

### **Список литературы**

1. Волкова В.Н., Денисов А.А. Методы организации сложных экспертиз: Учебное пособие. Изд. 4-е, переработанное и дополненное. СПб.: Издательство Политехн. университета, 2010. 128 с.
2. Волкова В.Н. Системный анализ информационных комплексов: Учеб. пособие. СПб.: Изд-во Политехн. ун-та, 2014. С.302–308.
3. A Guide to the Project management body of knowledge(PMBOK Guide). 6th Edition. Philadelphia, Pennsylvania, USA: Project Management Institute, 2017. 756 p.
4. Agile Practice Guide. Philadelphia, Pennsylvania, USA: Project Management Institute, 2017. 210 p.
5. Швабер К., Сазерленд Дж. Исчерпывающее руководство по Скраму: Правила игры. Ноябрь 2017. 26 с. URL: <http://brightmagazine.ru/wp-content/uploads/2019/01/2017-Scrum-Guide-Russian.pdf> (дата обращения 30.05.2020).

УДК 005; 65.015.11

doi:10.18720/SPBPU/2/id20-250

*Рудовская София Владимировна,*  
магистрант, бакалавр

## **АВТОМАТИЗАЦИЯ ПРИНЯТИЯ РЕШЕНИЙ ПО РАСПРЕДЕЛЕНИЮ СТАТЕЙ ПО РУБРИКАМ И ОЧЕРЕДНОСТИ ПРЕДСТАВЛЕНИЯ ИНФОРМАЦИИ НОВОСТНОГО САЙТА**

Санкт-Петербургский политехнический университет Петра Великого,  
Санкт-Петербург, Россия,  
[rudovskaya.sofia@yandex.ru](mailto:rudovskaya.sofia@yandex.ru)

*Аннотация.* В статье описаны методические основы создания информационного комплекса для автоматизации принятия решений о распределении статей по рубрикам, определении порядка представления информации, подбору визуального контента для сайта новостного издательства «Комсомольская правда».

Приведен сравнительный анализ существующей и предлагаемой систем методом экспертных оценок и методом, основанном информационном подходе А.А. Денисова.

**Ключевые слова:** новостной сайт, информационно-поисковая система, экспертные оценки, информационные оценки, системно-целевой подход, информационный комплекс, организация сложных экспертиз.

*Sofia V. Rugovskaya,*  
Master Student, BSc

## **AUTOMATION OF DECISION-MAKING ON THE DISTRIBUTION OF ARTICLES BY RUBRICS AND THE SEQUENCE OF SUBMITTING INFORMATION OF THE NEWS SITE**

Peter the Great St. Petersburg Polytechnic University,  
St. Petersburg, Russia,  
rudovskaya.sofia@yandex.ru

**Abstract.** The article describes the methodological foundations of creating an information complex for automating decision-making on the distribution of articles by category, determining the order of presentation of information, and selecting visual content for the website of the Komsomolskaya Pravda news publishing house. A comparative analysis of the existing and proposed systems by the method of expert assessments and the method based on the information approach of A.A. Denisov.

**Keywords:** news site, information and search system, expert evaluations, information estimations, system-target approach, information complex, complex expertise organization.

### **Введение**

Важную роль в деятельности средств массовой информации играют веб-сайты издательств. Успех сайта напрямую зависит не только от содержащейся на нём информации, но и от его структуризации, для чего используется категоризация статей и определенные правила, по которым определяется порядок представления статей. Также, на восприятие информации влияет дополнение статей визуальным контентом – фото и видео. Цель работы – разработка методических основ создания информационного комплекса для автоматизации принятия решений о распределении статей по рубрикам, определении порядка представления информации, подбору визуального контента. В качестве организации рассматривается издательский дом «Комсомольская правда».

### **1. Постановка задачи**

Определение системы, реализующее системно-целевой подход [1]:

$$S \stackrel{\text{def}}{=} \langle Z, STR, TECH, COND, N \rangle, \quad (1)$$

где  $N = \langle n_1, n_2, \dots, n_k \rangle$  – пользователи;  $Z = \langle z_1, z_2, \dots, z_i \rangle$  – новостные рубрики,  $STR = \langle str_1, str_2, \dots, str_m \rangle$  – структуры взаимосвязи  $Z$  с информационными ресурсами, где  $m = 2$ ,  $str_1$  – совокупность рубрик сайта с новостными

сообщениями,  $str_2$  – совокупность рубрик сайта с визуальными материалами;  $TECH = \langle tech_1, \dots, tech_n \rangle$  – совокупность технологий, используемых в издательстве и необходимых для реализации системы;  $COND = \langle cond_1, \dots, cond_s \rangle$  – условия существования системы, факторы, влияющие на её создание и функционирование.

Также есть множество  $MT = \langle mt_1, mt_2, \dots, mt_j \rangle$ , представляющее совокупность двух групп материалов, которые предполагается распределять по рубрикам – статьи и визуальные материалы (фото, видео).

Тогда задачу можно формализовать следующим образом:

$$Z \Psi MT, \quad (2)$$

где  $Z = \langle z_1, z_2, \dots, z_i \rangle$  – новостные рубрики,  $MT = \langle mt_1, mt_2, \dots, mt_j \rangle$  – множество материалов, подлежащих распределению по рубрикам,  $\Psi$  – сложный функционал, представляющий собой совокупность методов структуризации целей и функций, организации сложных экспертиз и других методик, позволяющих реализовать структуры.

## 2. Описание предлагаемого комплекса

Разрабатываемый информационный комплекс должен автоматизировать подбор рубрик и визуальных материалов для статьи, а также улучшить качество подачи контента на сайте путём определения порядка следования статей на основе рубрики и относительной важности статьи. Функционально процесс публикации статьи представлен на рисунке 1.



Рис. 1. Процесс публикации в разрабатываемом комплексе

Информационная структура разрабатываемого комплекса представлена в таблице 1.

**Информационная структура разрабатываемого комплекса**

<i>Страта 1. Пользовательская</i> Дежурный по сайту, штатные корреспонденты, веб-редактор, фотокорреспондент, главный редактор.
<i>Страта 2. Функциональная</i> Автоматический подбор рубрик, автоматический подбор визуального контента, определение важности статьи, добавление рубрик вручную, добавление визуального контента вручную, отправка статьи на сайт.
<i>Страта 3. Информационных ресурсов</i> Набор рубрик, архив фотографий, алгоритм взаимодействия статей и рубрик, алгоритм взаимодействия визуального материала и рубрик, алгоритм определения важности статей.
<i>Страта 4. Коммуникационная</i> Компьютеры, серверы

### **3. Описание структур взаимосвязи рубрик с новостными сообщениями и визуальными материалами**

Задачу построения структуры взаимосвязи рубрик со статьями и визуальными материалами можно представить как задачу информационного поиска. Информационно-поисковую систему можно описать по предложению А. И. Черного в виде совокупности четырех основных компонентов [5]:

$$IPS \stackrel{\text{def}}{=} \langle LS, D, TS, N \rangle, \quad (3)$$

где  $LS$  – логико-семантический аппарат;  $D$  – поисковый массив рубрик;  $TS$  – базы данных рубрик, статей, визуальных материалов, тегов, алгоритмы сопоставления рубрик и документов с тегами, алгоритмы сопоставления поисковых образов рубрик с поисковыми образами запросов, компьютеры, сервера и т. д.;  $N$  – дежурный по сайту, корреспонденты, веб-редактор, фотокорреспондент, главный редактор.

Для формирования поисковых образов рубрик можно совместить два типа систем индексирования в следующем порядке: сформировать архив распределенных по рубрикам статей, далее с помощью статистического и семантического анализа статей найти ключевые слова – теги – для каждой рубрики; проверку найденных тегов провести вручную экспертами, добавить упущенные ключевые слова. Подобным образом формируется поисковый образ статьи.

Так как для определенной статьи или фотографии может быть определено несколько рубрик, то для рассматриваемой системы выбирается критерий смыслового соответствия «на частичное вхождение» [4]. Планируемая структура взаимодействия элементов поисковой системы представлена на рисунке 2.

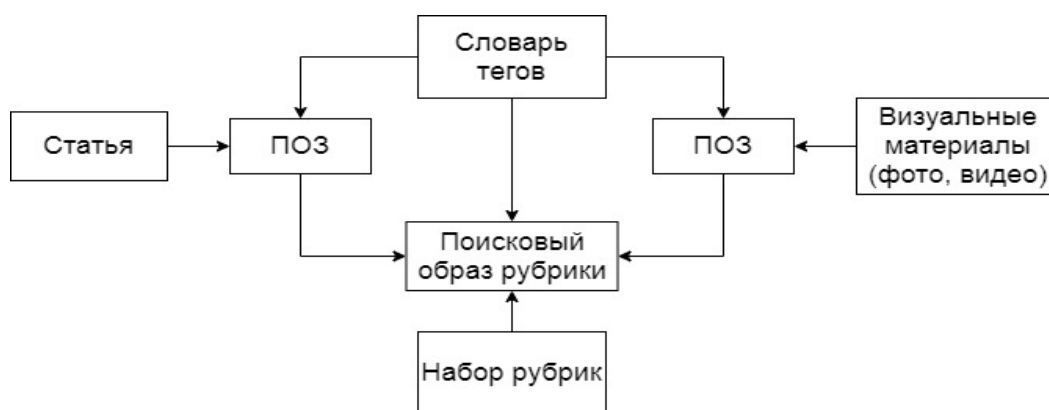


Рис. 2. Структура взаимодействия элементов поисковой системы рубрик

Подбор рубрики под соответствующую статью, фотографию или видео осуществляется следующим образом: для статьи или визуального материала формируется поисковый образ запроса; поисковый образ запроса сравнивается с поисковым образом рубрики; если определено частичное вхождение – запросу ставится в соответствие рубрика.

#### 4. Описание методики определения порядка следования статей на сайте

Определены веса рубрик: Политика – 0.9, Экономика – 0.8, Общество – 0.9, В Мире – 0.9, Спорт – 0.6, Звёзды – 0.6, Здоровье – 0.6, Недвижимость – 0.5, Наука – 0.5. У временных рубрик вес всегда будет равен единице.

Следует учитывать, что у статьи может быть несколько рубрик. Тогда оценка веса статьи по рубрикам можно определить по формуле 4:

$$VR = \frac{1}{n} \sum_{i=1}^n vr_i, \quad (4)$$

где  $n$  – количество рубрик статьи,  $vr_i$  – веса рубрик.

Кроме того, необходимо ранжирование внутри рубрик. Например, с помощью оценки важности, которую объективно может выставить только корреспондент.

Тогда итоговая оценка целесообразности новостной статьи представлена формулой 5.

$$VArt = VR * imp, \quad imp \in [0,1] \quad (5)$$

где  $VR$  – оценка веса статьи по рубрикам,  $imp$  – оценка важности статьи.

#### 5. Оценка целесообразности существующей и предлагаемой систем

Определена структура целей издательства при работе с сайтом, и на её основе для оценки целесообразности использовано два метода: метод экспертных оценок на иерархических структурах и метод, основанный на

информационном подходе А.А. Денисова [3]. Методы применены к существующей и разрабатываемой системе.

Так как существующая система уже используется в редакции, то вероятности использования  $q_i = 1$ , тогда соответствующая формула метода А.А. Денисова [2] примет вид:

$$H_i = -\log(1 - p_i'). \quad (6)$$

Итоговые оценки значимости вычисляются путём суммирования оценок отдельных функций. Оценки  $p_i'$  по методу представлены на рисунке 3.

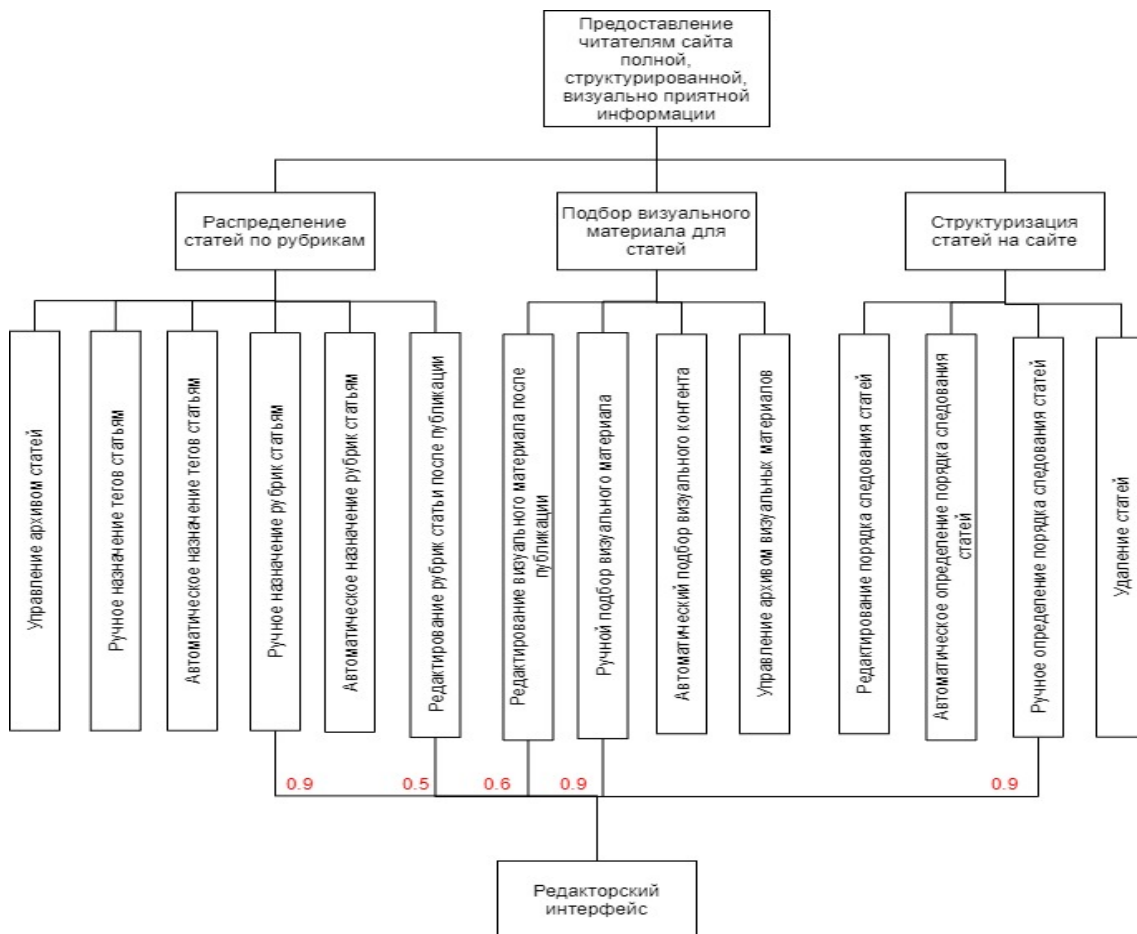


Рис. 3. Применение метода информационных оценок А.А. Денисова для существующей системы

Итоговая оценка значимости существующей системы:  $H_1 = 3,7$ .

Аналогичный анализ для разрабатываемой системы приведен на рисунке 4.

Итоговая оценка значимости предлагаемой системы:  $H_2 = 8,34$ .

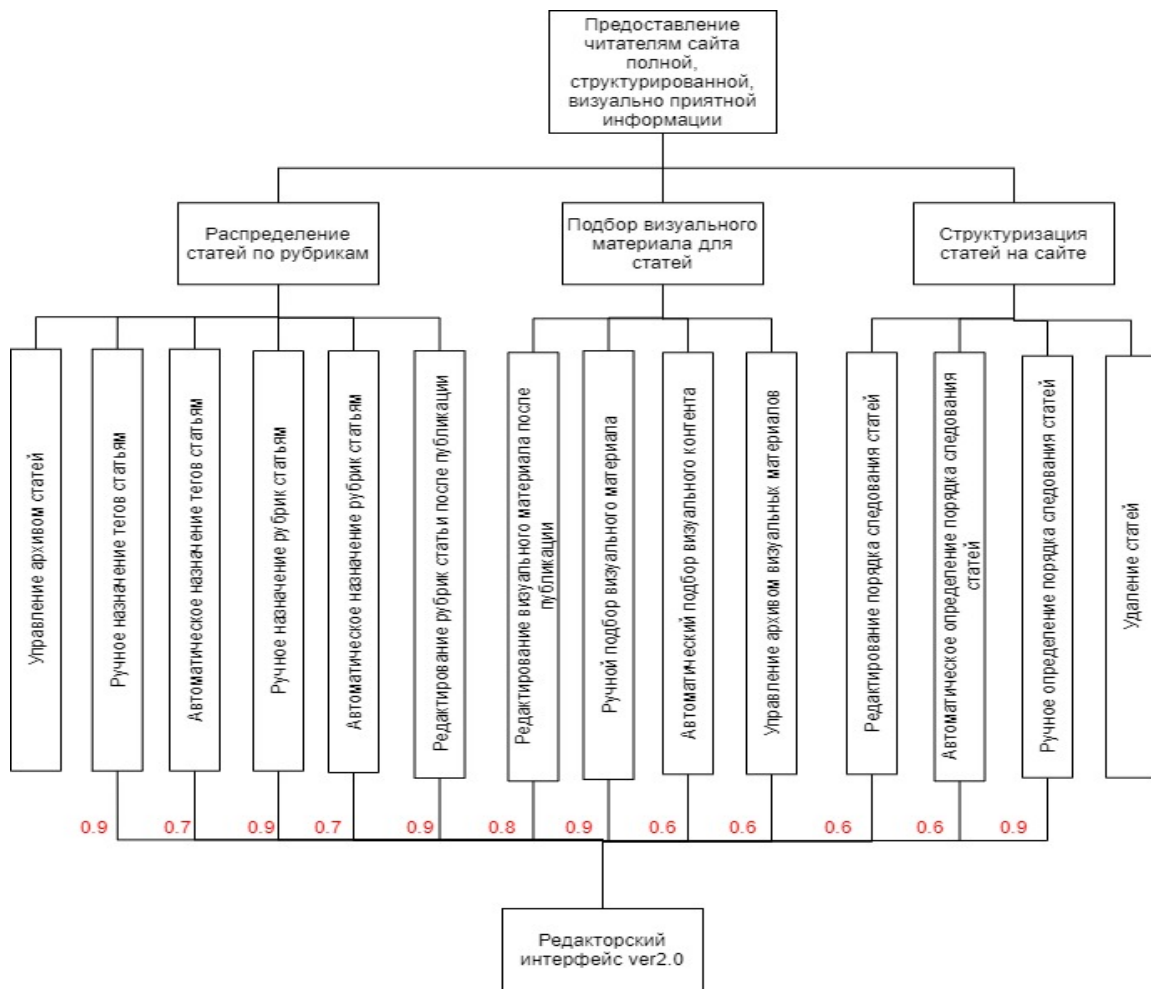


Рис. 4. Применение метода информационных оценок А.А. Денисова для разрабатываемой системы

Схема алгоритма реализации метода приведена на рисунке 5.

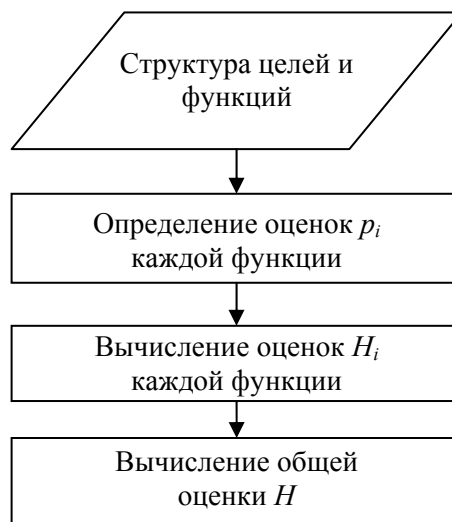


Рис. 5. Схема алгоритма реализации метода, основанного на информационном подходе А.А. Денисова

Подобным образом применен метод экспертных оценок, итоговые оценки: существующий интерфейс – 0,31; предлагаемый интерфейс – 0,9.

Оба метода показали существенное преимущество новой системы. Можно сказать, что существующая система реализует минимально необходимые функции, но внедрение новой системы может упростить и ускорить работу сотрудников по управлению сайтом издательства. Разрабатываемая система практически полностью покрывает потребности, это достигается как добавлением новых функций, так и путём совмещения в одной системе функций, которые в настоящий момент реализуются с помощью нескольких технологий.

### **Заключение**

В ходе работы были предложены улучшения существующей системы работы с сайтом, применяемой в редакции издания «Комсомольская правда», а именно: автоматическое распределение статей по рубрикам, подбор визуального контента к статье, определение порядка следования статей на сайте. Также, была разработана структура взаимосвязи статей и визуальных материалов с рубриками сайта, методика определения порядка следования статей. Далее произведена оценка существующей и предлагаемой системы с помощью метода экспертной оценки и метода организации сложных экспертиз, основанного на информационном подходе А.А. Денисова. Судя по оценкам, предлагаемая система может сэкономить время сотрудников, убрать воздействие некоторых человеческих факторов. Также, в ходе работы проводилась формализация задачи, определение методики выполнения работы, выбор методов оценки, структуризация целей и функций издания при работе с сайтом, описание существующей и предлагаемой систем.

### **Список литературы**

1. Волкова В.Н. Развитие определения системы // Матер. Международной научно-практической конференции «Системный анализ в проектировании и управлении»: сб. СПб.: Изд-во СПбГПУ, 2001. С. 12–14.
2. Волкова В.Н., Денисов А.А. Теория систем: Учебное пособие. М.: Высшая школа, 2006. 511 с.
3. Денисов А.А. Современные проблемы системного анализа: Учебник. 3-е изд., перераб. и дополн. СПб.: Изд-во Политехн. ун-та, 2009. 304 с.
4. Михайлов А.И., Черный А.И., Гиляревский Р.С. Основы информатики. М.: Наука, 1968. 756 с.
5. Черный А.И. Введение в теорию информационного поиска. М.: Наука, 1975. 298 с.