УДК 631.171:658.011.56

Меденников В.И.,
Сальников С.Г.
Федеральное государственное бюджетное научное учреждение «Всероссийский институт аграрных проблем и информатики имени А.А. Никонова»

Моделирование электронной торговой площадки на основе типизации сайтов сельскохозяйственных предприятий в ЕИПАЗ

Ключевые слова: электронная торговая площадка, Единое Интернет-пространство аграрных знаний, информатизация, интернет-технологии.

Реферат. В работе рассматриваются возможности развития электронной торговли на основе типизации сайтов сельскохозяйственных предприятий в рамках Единого Интернет-Пространства Аграрных Знаний.

Abstract. This paper examines the opportunities of development of electronic commerce based on the typification of agricultural enterprise sites within the Common Internet-Space of Agrarian Studies

Опыт наиболее развитых стран показывает, что обеспечить переход к инновационному развитию сельского хозяйства можно только, используя современные информационно-сетевые технологии Интернета.

Хаотичное, неуправляемое развитие Интернет-технологий со стороны государства делает процесс их внедрения более длительным и приводит к значительному перерасходу ресурсов. Даже те незначительные ресурсы, что выделяются на развитие Интернет-технологий, используются крайне неэффективно.

Согласно исследованиям, проведенным ВИАПИ, сайты научных организаций, как и сайты российских госструктур, образовательных учреждений малоинформативны и практически не предоставляют никаких онлайновых услуг. Даже если сайт насыщен полезной информацией, то она плохо структурирована, нет понятной пользователю системы навигации, отсутствует возможность обмена информацией между сайтами и т.д.

Ситуация с Интернет-технологиями в сельскохозяйственных предприятиях еще более удручающая. Лишь менее 5% предприятий имеют сайты, как правило, очень низкого качества.

Конечно, сейчас делается много попыток интеграции информации в Интернете. Например, в БД «Спарк» Интерфакса [1] собирается информация по всем предприятиям России, стран СНГ по 80 показателям. Информация покупается из ГНИ, статистических органов, таможни и т.д. Данная работа требует больших усилий по сбору, рубрикации, унификации данных. Всем известны БД «Price» Рамблера и «Маркет» Яндекса. Так, для актуализации БД «Price» свыше 70 человек ежедневно собирают по всем торговым сайтам нужную информацию, затем обрабатывают ее и помещают в БД.

Наиболее бурно в мире развивается Интернет-торговля. В сельском хозяйстве РФ также стихийно создаются сайты, отражающие, в той или иной степени, автоматизацию информационных процессов реализации продукции и услуг сельскохозяйственных предприятий в Интернете.

Например, по прогнозам ВШЭ к концу десятилетия Интернет будет обслуживать более 50% российской экономики.

Актуальность грамотного проектирования электронных торговых площадок продиктована тем, что наибольший эффект для товаропроизводителей дают Интернет-технологии в сфере электронной торговли.

Анализ состояния процессов реализации продукции и услуг сельскохозяйственных предприятий в Интернете позволяет выделить следующие последовательные уровни автоматизации данных процессов:

Создание информационного поля для реализации поиска продукции в виде общей неструктурированной доски объявлений;

Структурирование объявлений по видам (категории) товаров и услуг с использованием произвольно составленных классификаторов товаров и услуг;

Автоматизация поиска оптимального торгового партнера по заданному показателю (критерию оптимальности);

В дополнение к п.3 осуществляется учет общих затрат на приобретение товара, включая транспорт, налоги и др.;

Автоматизация информационных процессов всех торговых операций;

Полная автоматизация электронной торговли с юридическим оформлением торговых операций.

Можно выделить **два варианта реализации этих процессов**.

Первый вариант - создается организационная структура, которая, подобно Рамблеру или Яндексу, собирает в одной базе данных с определенной периодичностью всю информацию по всем сайтам, имеющим спрос и предложение товаров и услуг в области сельского хозяйства. На данном интегрированном сайте создается информационная система, реализующая тот или иной уровень автоматизации электронной торговли.

Второй вариант – разрабатывается типовой сайт сельскохозяйственного предприятия с размещением у одного провайдера под управлением мощной СУБД на основе единых классификаторов, например, ГРНТИ и ОКП. Данный подход дает огромные преимущества. Это возможность различной аналитической обработки информации, в частности, построение различных рейтингов, выборки, группировки, расчеты индексов цитирования, выявление плагиата, а также сведение напрямую продавцов и покупателей с расчетом транспортного плеча и оптимизацией издержек, проведение целенаправленной миграции трудовых ресурсов, сдача в лизинг программного обеспечения, проведение ценового мониторинга и т.д. В предложенном варианте информация автоматически находится в общей БД в соответствующих разделах без дополнительных затрат на ее сбор и актуализацию.

Структура типового сайта приведена в [2]. Раздел, посвященный электронной торговле, выглядит следующим образом:

- Дата публикации;
- Наименование продукта (услуги);
- Вид (тип, группу, сорт) продукта (услуги);
- Количество продукта;
- Цена продукта (услуги);
- Тара;
- Условия поставки (дополнительная информация).

Использование типового сайта сельскохозяйственных предприятий на основе мощной СУБД и единых классификаторов, таким образом, позволит осуществить автоматизацию торговых операций 4 и выше уровней с поиском наиболее выгодного партнера. В плановой экономике можно было бы решать транспортную задачу минимизации всех суммарных затрат, указанных выше, для всех поставщиков и производителей. На данный момент реально минимизировать такие затраты лишь для конкретного пользователя системы. Алгоритм решения данной задачи с точки зрения конкретного потребителя в рамках типизации сайтов сельскохозяйственных предприятий описывается следующим образом.

Введем обозначения.

v_{ik} - объем потребности i - го потребителя в k - ом продукте, $i \in$; $k \in$;

w_{jk} - объем наличия k - го продукта у j - го производителя, $j \in$;

p_{jk} - цена k - го продукта у j - го производителя;

m_{ijr} - расценки на транспортировку из пункта нахождения j - го производителя в пункт нахождения i - го потребителя r-м видом транспорта;

$v^1_{ijk} = v_{ik}$, если $v_{ik} \le w_{jk}$, и $v^1_{ijk} = w_{jk}$ иначе;

$c^1_{ijrk} = v^1_{ijk} \text{x}\, m_{ijr}$ - расходы на транспортировку k - го продукта от j - го производителя i - му потребителю r-м видом транспорта;

$c^2_{ijk} = v^1_{ijk} \text{x}\, p_{jk}$ - стоимость k - го продукта, приобретаемого i - м потребителем у j - го производителя;

$c_{ijrk} = c^1_{ijrk} + c^2_{ijk}$ - общая стоимость k - го продукта, приобретаемого i - м потребителем у j - го производителя.

Тогда,

шаг 1: введем множество $J^* \subset$, $J^* =$;

шаг 2: решается задача $c_{ijrk} = \rightarrow \text{in}$ по переменным r и j при ограничениях $v_{ik} \le \sum_j w_{jk}$, $j \in$ / J^*. Найденное решение j^* включается во множество J^*;

шаг 3: выполняется шаг 2 до тех пор, пока не будут найдены все производители, удовлетворяющие критерию шага 2, то есть способные поставить необходимое количество нужного продукта конкретному потребителю. При этом они будут ранжированы в порядке возрастания суммарной стоимости k - го продукта.

Дальнейшее взаимодействие потребителю и производителя зависят от уровня автоматизации торговых операций, указанных выше.

Сведения о расстояниях, видах транспорта и тарифах предоставляют транспортные учреждения (соответствующие базы данных широко представлены в Интернете). Номера или названия пунктов выбираются по адресу учреждения, указанного в разделах типового сайта.

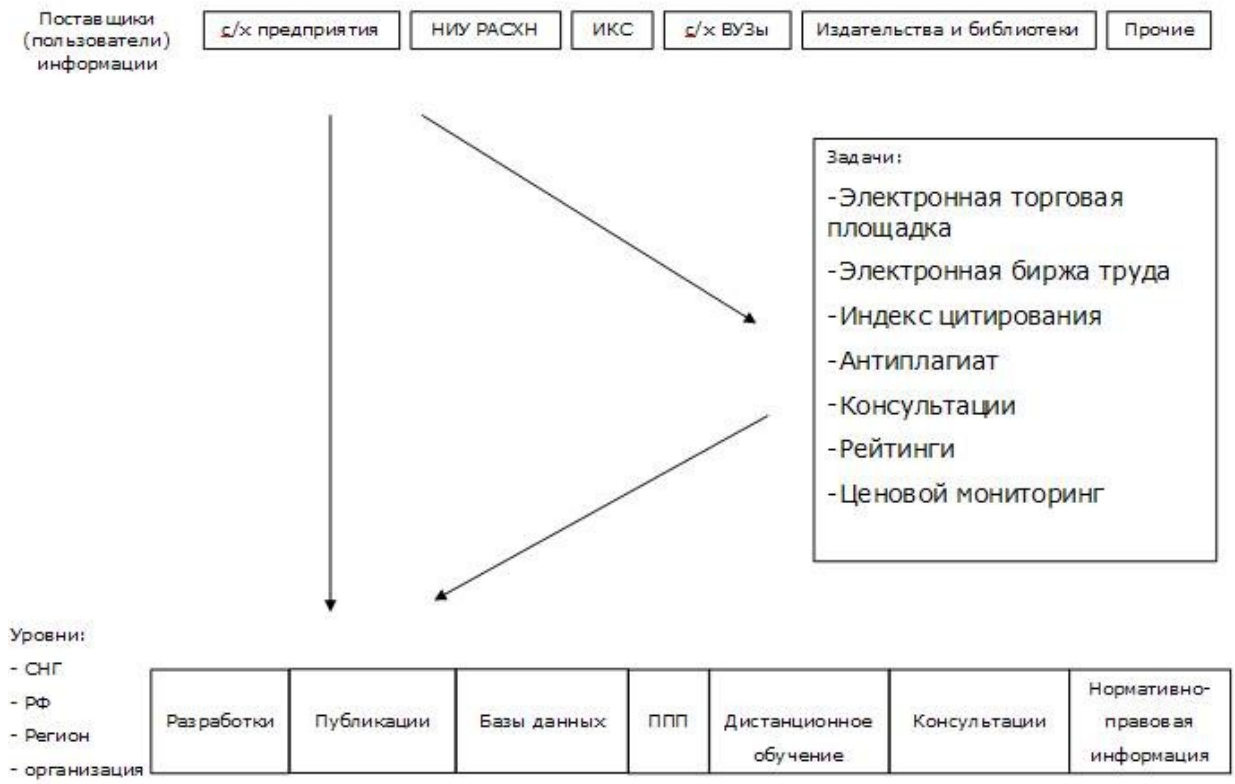


Рис. 1 Структура ЕИПАЗ

Процесс создания электронной торговой площадки на базе типового сайта сельскохозяйственных предприятий должен быть интегрирован в единое Интернет-пространства аграрных знаний (ЕИПАЗ), структура которого представлена на рисунке 1, для чего необходимо провести моделирование возможных путей интеграции разработок, публикаций, консультационной деятельности, нормативно-правовой информации, дистанционного обучения, пакетов прикладных программ, баз данных, разработанных агропромышленными НИИ, ВУЗами сельскохозяйственного профиля, НИИ РАН и другими организациями, занимающимися сельскохозяйственной тематикой, с единых позиций в целях облегчения поиска потребителями нужной информации с размещением у одного провайдера под управлением мощной СУБД на основе единых классификаторов, например, ГРНТИ и ОКП.

На рисунке 2 представлены возможные варианты такой интеграции.

В 1-м варианте предполагается перенос на единую базу данных у единого провайдера каталогов информационных массивов. При удачном поиске нужной информации в каталоге пользователь отсылается на сайт, хранящий полнотекстовую, либо более подробную информацию. Второй вариант отражает ситуацию, когда вся информация перенесена к единому провайдеру. В третьем варианте часть информации перенесена в виде каталогов, а часть полностью. Например, переносится такие виды представления данных, как публикации, разработки, нормативно-правовая информация, т.е. те виды, поиск по которым требует быстрого перебора всех сайтов.

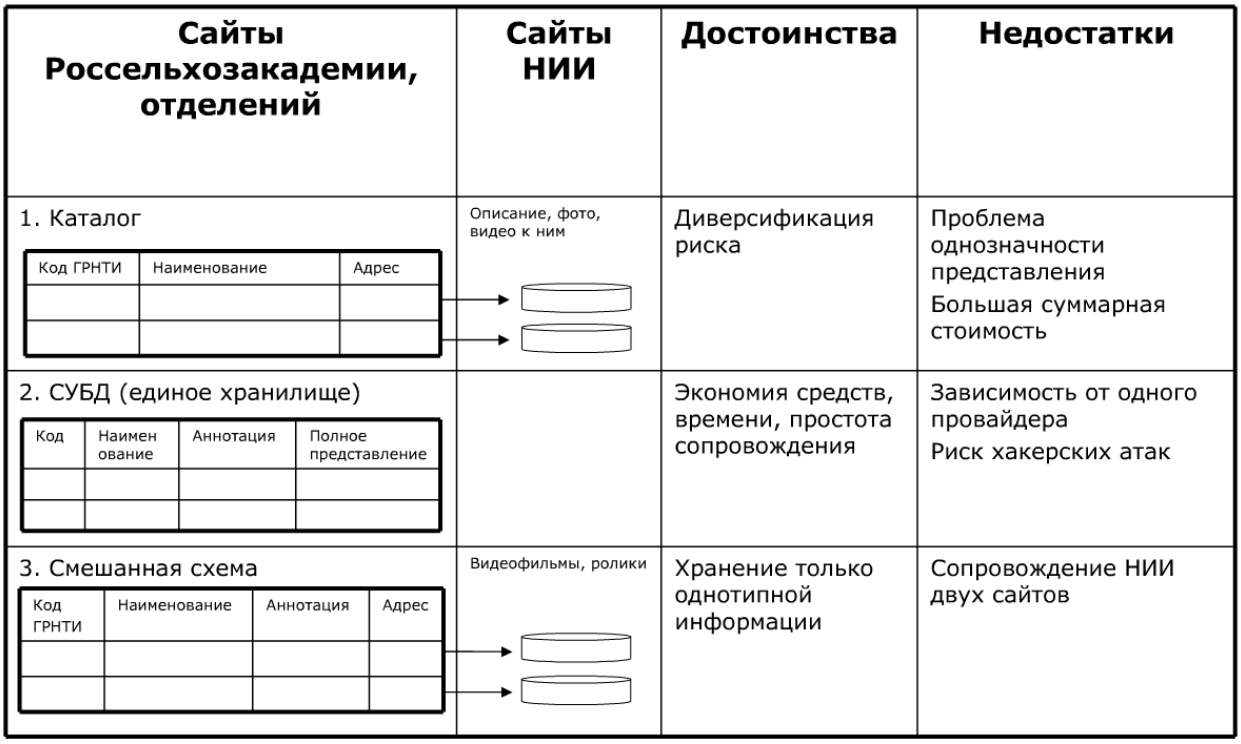


Рис. 2 - Варианты интеграции информации в ЕИПАЗ

Качество обслуживания провайдером пользователей при этом обычно измеряется несколькими параметрами: надежностью сети, временными задержками при передаче информации, статистическими характеристиками задержек, пропускной способностью. Поскольку мы хотим выяснить глобальные характеристики сети при переносе информации большим количеством владельцев ее к одному провайдеру, не обладая конкретными объемами этой информации (на сайтах выложена лишь незначительная часть ее в плохо структурированном виде), то мы будем моделировать процессы на достаточно большом интервале времени, например, месяц. На данном промежутке актуальность имеет лишь пропускная способность сети.

Описание модели ЕИПАЗ

1 Константы и параметры.

Под «единицей времени» всюду ниже понимается месяц

i - код вида информации (тексты, изображения, видео и др.), $i \in$;

j - код группы организаций-носителей информации (ВУЗы, НИУ и др.), $j \in$;

m - код провайдера, использующего Битрикс, $m \in$;

n - код вида представления данных, $n \in$;

k - код конкретной организации-носителя информации; $k \in$;

l - код формы хранения информации, $l \in$;

d_{im} - существующая нагрузка m - го провайдера на i - ый вид информации (в Мбайтах);

V_{ijkl} - объём i - го вида информации для l-ой формы хранения информации k-ой организации j - ой группы организаций (в Мбайтах);

D_{im} - пропускная способность m - го провайдера i - го вида информации (в Мбайтах);

z^1_{ijmkl} - средние удельные затраты по переносу к провайдеру, работающему на Битрикс, единицы i - го вида информации для l-ой формы хранения информации j - ой группы m - го провайдера k-ой организации (в рублях/Мбайт);

z^2_m - общие затраты в единицу времени на сопровождение сайта у m - го провайдера (в рублях);

z^3_{jk} - общие затраты в единицу времени на сопровождение сайтов j - ой группы организаций при хранении информации у своего провайдера k-ой организации (в рублях);

P^1_i - среднее количество обращений (количество посетителей) к i - му виду информации в единицу времени;

P^2_{il} - среднее число просмотров страниц i - го вида информации для l-ой формы хранения информации;

P^2_{ijkl} - среднее число просмотров страниц j - ой группы организаций i - го вида информации для l-ой формы хранения информации k-ой организации;

$P^2_{il} = s \cdot \sum_{j,k} P^2_{ijkl}$, где s - коэффициент увеличения за счёт интеграции (у нас при переходе на типовой сайт считается, что $s = 1{,}5$);

P^3_m - средний размер страницы сайта у m - го провайдера (в Мбайтах);

C^0 - средства, выделенные в единицу времени на перенос информации к одному из провайдеров m (в рублях);

В дальнейшем будем считать, что у любого провайдера, использующего Битрикс, вся информация будет храниться в унифицированном виде, аналогичном таковой на Портале Россельхозакадемии.

Тогда введём ещё группу параметров:

b_{il} - средний размер в единицу времени передаваемого файла i - го вида информации l - ой формы хранения информации (в Мбайтах);

g_{il} - среднее количество обращений для i - го вида информации l - ой формы хранения информации у любого провайдера, использующего Битрикс;

r_{jknl} - количество видов представления данных n - го вида представления информации для j - ой группы l-ой формы хранения информации k-ой организации у своего провайдера;

a_{iknl} - среднее количество обращений (количество посетителей) n - го вида представления информации для i - го вида информации l-ой формы хранения информации у своего провайдера;

v_{inl} - индекс, отражающий наличие i - го вида информации l-ой формы хранения информации у n - го вида представления данных;

$$v_{inl} = \begin{cases} 1 - \text{если имеется } i\text{-ый вид информации } l\text{-ой формы у } n\text{-го вида представления данных;} \\ 0 - \text{иначе} \end{cases}$$

Тогда:

$$g_{il} = s \cdot \sum_{n,k} v_{inl} \cdot a_{iknl};$$

$$V_{ijkl} = b_{il} \cdot \sum_{n} r_{jknl};$$

$$P_i^1 = \sum_{l} g_{il};$$

2 Переменные.

x_{ijmkl} - увеличение нагрузки на m - го провайдера за счёт размещения у него i - го вида информации для l-ой формы хранения информации j - ой группы k-ой организации (в Мбайтах);

$$y_{ijmkl} = \begin{cases} 1 - \text{если для k-ой организации j-ая группа хранит i-ый вид информации} \\ \quad \text{в l-ой форме хранения информации у m-го провайдера} \\ 0 - \text{иначе} \end{cases};$$

3 Уравнения модели.

$d_{im} + \sum_{l,j,k} x_{ijmkl} \le D_{im}$ - ограничения на пропускные способности m - го провайдера для i - го вида информации; (1)

$x_{ijmkl} = (P_i^1 \cdot P_{il}^2 \cdot P_m^3 + r_{il} \cdot b_{il}) \cdot y_{ijmkl}$ - балансовое равенство для добавочной нагрузки; (2)

$\sum_{m} y_{ijmkl} \le 1$ - ограничения на то, что вся информация может храниться только у одного из провайдеров; (3)

m

Наконец, для затрат по переносу информации к провайдеру работающему на Битрикс

$$C^1 = \sum_{i,j,m,l,k} z^1_{ijmkl} \cdot V_{ijkl} \cdot y_{ijmkl}$$

имеем неравенство:

$C^1 \leq$ - ограничения по затратам на перенос (4)

4 Оптимизационные критерии.

Таковых у нас в модели два:

$$w = \sum_{i,j,m,l,k} V_{ijkl} \cdot y_{ijmkl} \rightarrow \max$$

- максимизация объёмов переноса информации к Битрикс-провайдерам (5)

или

$C^2 \rightarrow \min$ - минимизация затрат на сопровождение у Битрикс- провайдеров (6)

где

$$C^2 = T \cdot \left(\frac{1}{I \cdot J \cdot L} \sum_{i,j,m,l,k} z_m \cdot y_{ijmkl} + \frac{1}{I \cdot M \cdot L} \sum_{i,j,m,l,k} z_{jk} \cdot (1 - y_{ijmkl})\right)$$

T - заданный период эксплуатации системы (в месяцах, в силу нашего первоначального выбора месяца в качестве единицы времени);

Для нахождения решения можно воспользоваться стандартными методами теории графов [3]. Однако на данном этапе вполне можно, учитывая, что провайдеров, имеющих инструментальное средство для ведения интегрированных порталов «Битрикс», не очень много, применить двухэтапную процедуру оптимизации, заключающуюся в следующем. На первом этапе последовательно решается задача оптимизации для каждого провайдера, фиксируя конкретное значение m. На втором этапе выбираем наиболее приемлемого провайдера с точки зрения выбранного критерия.

Для сценарных расчётов в модели будем рассматривать 3 базовых сценария по объёмам информации – текущий объём, 5-летний объём (информация за 5 последних лет), полный объём (информация за все годы) – и 2 базовых сценария по численности посетителей сайтов – текущая численность и полная численность (максимально прогнозируемое количество).

Для оценки объемов информации за 2 последних базовых сценария воспользуемся методикой технико-экономического обоснования (ТЭО) работ по созданию СНТИ [4]. Так, объем печатной продукции Россельхозакадемии для 2-го сценария составляет 455850 печатных листов, для 3-го – 15954750 печатных листов.

В таблице 1 приведены текущие объемы имеющихся данных на сайтах различных групп носителей информации для 1-го базового сценария.

Будем считать, что количество форм хранения информации для всех видов для 2-го и 3-го сценариев совпадает, то есть все публикации будут на 100% перенесены на машинные носители. Примем, что объемы всех видов информации для всех групп организаций для 2-го и 3-го сценариев возрастут пропорционально публикациям в Россельхозакадемии для этих сценариев.

Таблица 1 - Объёмы информации для 1-го базового сценария

		Виды представления данных						
Группа	Форма хранения	Разработки	Публикации	БД	Дист.образ.	ППП	Конс.	НПИ
НИУ Россельхозакадемии	Каталожная	1158	1172	3	2	3	7	338
НИУ РАН	Каталожная	34	165	-	-	-	-	-
НИУ Медицинской академии	Каталожная	-	-	-	-	-	-	-
ВУЗы сельхозпрофиля	Каталожная	1038	11783	4	13	32	236	-
Издательства аграрной тематики	Каталожная	-	9166	-	-	-	-	-
НИУ стран СНГ	Каталожная	191	65	-	-	-	-	1
НИУ Россельхозакадемии	Полнотекстовая	2677	2372	46	1	107	4	175
НИУ РАН	Полнотекстовая	-	184	-	-	-	-	-
НИУ Медицинской академии	Полнотекстовая	-	-	-	-	-	-	-
ВУЗы сельхозпрофиля	Полнотекстовая	-	-	-	-	-	-	-
Издательства аграрной тематики	Полнотекстовая	-	7084	-	-	-	-	-
НИУ стран СНГ	Полнотекстовая	267	51	-	-	-	-	4

Численность посетителей ЕИПАЗ оценим экспертно на основании анализа различных оценок, присутствующих в Интернете. Кроме того, будем считать их количество для 2-го и 3-го сценариев одинаковым.

Тогда в таблицах 2 и 3 представлены объёмы информации для 2-го и 3-го базовых сценариев.

Таблица 2 - Объёмы информации для 2-го базового сценария

		Виды представления данных						
Группа	Форма хранения	Разработки	Публикации	БД	Дист. обр.	ППП	Конс.дии	НПИ
НИУ Россельхозакадемии	Каталожная	357388	364682	7293	4886	21879	2125	3756
НИУ РАН	Каталожная	5082	25410	-	-	-	78	-
НИУ Медицинской академии	Каталожная	939	1100	-	-	-	-	-
ВУЗы сельхозпрофиля	Каталожная	159683	1814582	8754	26262	24915	9788	788
Издательства агр. тематики	Каталожная	-	1411564	-	-	-	-	-
НИУ стран СНГ	Каталожная	29028	10010	-	-	-	-	1899
НИУ Россельхозакадемии	Полнотекстовая	357388	364682	7293	4886	21879	2125	3756
НИУ РАН	Полнотекстовая	5082	25410	-	-	-	78	-
НИУ Медицинской академии	Полнотекстовая	939	1100	-	-	-	-	-
ВУЗы сельхозпрофиля	Полнотекстовая	159683	1814582	8754	26262	24915	9788	788
Издательства агр. тематики	Полнотекстовая	-	1411564	-	-	-	-	-
НИУ стран СНГ	Полнотекстовая	29028	10010	-	-	-	-	1899

Таблица 3 - Объёмы информации для 3-го базового сценария

		Виды представления данных						
Группа	Форма хранения	Разработки	Публикации	БД	Дист. обр.	ППП	Конс.	НПИ
НИУ Россельхозакадемии	Каталожная	2501718	2552774	7293	4886	21879	2125	3756
НИУ РАН	Каталожная	35574	177870	-	-	-	78	-
НИУ РАМН	Каталожная	6574	7700	-	-	-	-	-
ВУЗы с/х профиля	Каталожная	1117782	12702074	8754	26262	24915	9788	788
Изд-ва агр. тематики	Каталожная	-	9880948	-	-	-	-	-
НИУ стран СНГ	Каталожная	203200	70070	-	-	-	-	1899
НИУ Россельхозакадемии	Полнотекстовая	2501718	2552774	7293	4886	21879	2125	3756
НИУ РАН	Полнотекстовая	35574	177870	-	-	-	78	-
НИУ РАМН	Полнотекстовая	6574	7700	-	-	-	-	-
ВУЗы сельхозпрофиля	Полнотекстовая	1117782	12702074	8754	26262	24915	9788	788
Издательства аграрной тематики	Полнотекстовая	-	9880948	-	-	-	-	-
НИУ стран СНГ	Полнотекстовая	203200	70070	-	-	-	-	1899

Полный сценарий по посетителям предполагает в качестве посетителей следующие типы и прогнозируемые количества (Таблица 4):

Таблица 4 - Прогнозируемые количества посетителей сайтов

Типы посетителей	Количество (в месяц)
Фермеры	200 000
Сотрудники с/х предприятий	1 000 000
Студенты	30 000 000
Управленческие работники	200 000
Научные сотрудники	1 200 000
Прочие	32 600 000
ВСЕГО	65 200 000

Остальные параметры модели являются усредненными данными по сайтам организаций различных их групп и провайдеров, работающих с Битрикс.

Основные выводы из представленных выше результатов можно сделать следующие.

1. Технические и программные возможности провайдеров, работающих на Битрикс, в настоящее время позволяют перенести, по крайней мере, всю информацию, все знания, накопленные за последние 5 лет, к одному из них в рамках выделенной пропускной способности провайдера.

2. Поскольку большинство научных знаний еще только предстоит оцифровать, на что потребуются большие средства и время, а каналы связи и соответствующие программные средства совершенствуются, в том числе и в сторону увеличения мощности, год от года, то нет видимых технических ограничений на перенос всех знаний, накопленных научными организациями в области сельского хозяйства, к одному провайдеру. Это позволит значительно сократить затраты на сопровождение как сайтов, так и программного обеспечение.

3. Из упомянутой выше методики [4] также следует, что самостоятельное преобразование в электронный вид всей научной продукции потребует значительных усилий и средств. Например, самостоятельное преобразование в электронный вид всех публикаций только Россельхозакадемии, как это делается на Западе, потребует свыше 450 млн. рублей. Поскольку количество публикаций в ВУЗах сравнимо с их количеством в Россельхозакадемии, то общие затраты превысят 900 млн. рублей. Затраты на разработку специализированных баз данных и пакетов прикладных программ составят около 200 млн. рублей [5]), Только объединив усилия, (например, у издательств имеются в электронном виде почти все публикации за последние годы) можно добиться существенного сокращения затрат на эту работу. При этом экономия на оцифровке публикаций составит около 900 млн. рублей и экономия на сопровождении сайтов превысит 2 млн. рублей в год.

4. Результаты моделирования вселяют уверенность в возможность разработки типового сайта сельскохозяйственного предприятия, информационно-консультационного центра с включением в них раздела «Научное сопровождение агропромышленного производства», «Документооборот», «Статистика», «Электронная площадка» для облегчения доступа к информационным Интернет-ресурсам, передачи его бесплатно им и размещения уже наполненных сайтов к одному провайдеру. Это позволит осуществлять возможность различной аналитической обработки информации, в частности, строить различные рейтинги, выборки, группировки, а также рассчитывать индексы цитирования.

5. Реализация портала ЕИПАЗ в таком виде позволит шагнуть сельскому хозяйству действительно в информационное общество.

ЛИТЕРАТУРА

1 СПАРК – Система профессионального анализа рынков и компаний [Электронный ресурс]. 2011. URL: http://spark.interfax.ru/Front/index.aspx.

2 Отчет о НИР «Анализ состояния и объемы информационных ресурсов сельскохозяйственных предприятий в Интернет-пространстве». -ВИАПИ РАСХН. 2011.

3 Кульба В.В., Ковалевский С.С., Косяченко С.А.,Сиротюк В.О.. Теоретические основы проектирования оптимальных структур распределенных баз данных.-М.: Синтег, 1999.

4 Отчет о НИР «Провести анализ состояния и тенденций развития систем информатизации аграрной науки на основе результатов мониторинга». -ВИАПИ Россельхозакадемии. 2006.

5 Отчет о НИР «Разработать проект программы развития информатизации аграрной науки». -ВИАПИ Россельхозакадемии. 2005.