

УДК 004.777:025.2

Опыт интеграции
информационных
систем ЦНБ УРО РАН
и информационных потоков
из внешних источников для
задач экспертной системы
комплектования

Л.Г. Горбич

Научный сотрудник,
E-mail: glg@cbibl.uran.ru

А.С. Павлова

Младший научный сотрудник, главный
библиотекарь,
E-mail: ocompl@cbibl.uran.ru

С.Г. Шмелева

Ведущий инженер-программист
E-mail: shsg261@rambler.ru
Центральная научная библиотека
Уральского отделения Российской
академии наук (ЦНБ УрО РАН), Россия,
Екатеринбург

Аннотация. В статье рассматриваются вопросы обмена данными как между информационными системами внутри библиотеки, так и с внешними информационными источниками на примере задач Экспертной системы комплектования. Представлено решение одной из типичных задач этого направления. Проанализированы достоинства и недостатки такого решения.

Ключевые слова. Библиотечная информационная система, экспертная система комплектования.

Experience of Integration
of the Information Systems
of the CSL UB RAS and
Information Flows from
External Sources for the Tasks
of the Expert System
of Library Acquisition

Gorbich L.G.

E-mail: glg@cbibl.uran.ru

Pavlova A. S.

E-mail: ocompl@cbibl.uran.ru

Shmeleva S.G.

Central scientific library of the Ural branch of
the Russian Academy of Sciences (CNS RAS),
Yekaterinburg, Russia,
E-mail: shsg261@rambler.ru

Abstract. The article discusses the issues of data exchange between information systems within the library and with external information sources for the tasks of the Expert System of Acquisition. It presents a solution of one of the typical problems in this direction and analyzes the advantages and disadvantages of this solution.

Keywords. library information system, expert system of library acquisition.

DOI: 10.31432/1994-2443-2018-13-4-27-31

Проблема совместного использования нескольких информационных систем и их связи с внешними источниками данных при решении актуальных задач организации возникает постоянно и закономерно. Каждый раз решение этой проблемы требует неформального, творческого подхода и оказывается существенно зависящим от большого количества разнообразных влияющих факторов. Рассмотрим, для начала, общие аспекты интеграции информационных систем в библиотеках, перейдя, в даль-

нейшем, к описанию конкретных информационных систем ЦНБ УрО РАН и отдельных задач, возникающих в процессе ее деятельности.

Информационные потоки из внешних источников являются весьма вариативными как во времени, так и по их структуре и используемым форматам представления данных. С другой стороны, собственные информационные системы, используемые в научных библиотеках и других академических институтах, хотя обычно и связаны информа-

ционно между собой, но эта связь, чаще всего, не является ни жестко стандартизированной по форматам, ни однозначно, пусть и по своему собственному формату, раз и навсегда установленной. Это обусловлено несколькими объективными причинами: во-первых, ни одна даже достаточно сложная и разветвленная система не покрывает всех нужд библиотеки, и приходится использовать несколько информационных систем, следовательно, и промежуточные форматы данных; во-вторых, каждая из этих систем, в силу своей специализации, оперирует со своим форматом и структурой базы данных и собственными интерфейсами взаимодействия с другими системами; в-третьих, набор используемых информационных систем меняется, модифицируются сами системы, поскольку изменяются требования и структура информационного взаимодействия внутри эксплуатирующей организации; в-четвертых, изменяются информационные потоки из внешних источников, что требует адаптации информационных систем. Отметим, что можно наблюдать и положительные, синергетические эффекты от интеграции систем, когда связанный комплекс систем позволяет решать задачи, решение которых невозможно в рамках каждой отдельной системы. Все это приводит к тому, что проблема обмена данными, как между информационными системами организации, так и с внешними системами всегда является актуальной.

Интерактивная система оценок (Экспертная система комплектования) введена в процесс информационного обеспечения научных исследований УрО РАН в 2008 г. [1]. Экспертная оценка осуществлялась с помощью системы экспертизы, имевшей свои форматы обмена данными (разработка БЕН РАН). Разработка и внедрение системы «Web-кабинет ученого», предназначенной для дистанционного интерактивного обслуживания пользователей-читателей библиотеки ЦНБ УрО РАН через сеть Интернет [2, 5], началась в 2013 г. В 2016 г. был разработан и вошел в состав системы «Web-кабинет учёного» — модуль, обеспечивающий экспертную оценку пользователями книг научной направленности с целью определения потребности в их закупке, который заменил систему, использовавшуюся до этого.

Рассмотрим особенности Экспертной системы комплектования в составе системы «Web-кабинета учёного». В структуре «Web-кабинета учёного» Экспертная система комплектования, по сравнению с предшествующей реализацией, имеет расширенный функционал, как для пользователей-учёных, так и для специалистов-комплектователей. В разделе «Книги» есть доступ к дополнительным вкладкам: «Изданные», «Экспертиза» и «Поступившие» — для экспертов и пользователей; а также «Предобра-

ботка», «Постобработка», «Редактирование», «Комплектование», «Отчеты по комплектованию» — для комплектователей.

Раздел «Изданные» носит ознакомительный характер и предоставляет информацию о научных изданиях, вышедших на территории России. В разделе «Экспертиза» учёные-эксперты могут оценить предложенные им издания, а в разделе «Поступившие» просмотреть, какие из заказанных изданий уже поступили в ЦНБ УрО РАН. Имеется функция персональных оповещений экспертов об изданиях, поступивших по их рекомендации. Разделы для комплектователей соответствуют последовательности стадий работы с наборами изданий в Экспертной системе.

Таким образом, в процессе комплектования Экспертная система выполняет следующие основные функции:

- информационная — информирование о документопотоке научных изданий в области научных исследований УрО РАН;
- сигнальная — информирование о выходе в свет изданий институтов и ученых УрО РАН;
- ориентирующая — отбор изданий по тематике научных исследований УрО РАН;
- оценочная — оценка изданий на соответствие тематике научных исследований УрО РАН;
- статистическая — предоставление статистических данных об информационных потребностях ученых УрО РАН в научных книжных изданиях.

До марта 2018 г. информация о вышедших в России книгах поступала из централизованного источника (Российской книжной палаты) в виде архивированных файлов формата RUSMARC (ГОСТ 7.1-2003), после чего она загружалась в систему, обрабатывалась, и формировались пакеты записей, выставляемые в дальнейшем на экспертизу. Следует отметить, что, хотя такой путь ввода информации в систему и обладал определенными преимуществами: полнота охвата, редкость дублирования записей, структурированность и стандартный формат; но за годы работы с Экспертной системой комплектования выявился ряд недостатков. Например, информация была явно избыточной: по данным мониторинга эффективности работы Экспертной системы на протяжении нескольких лет на этапе первичного редактирования отсеивалось до 90% получаемой библиографической информации [3, 4]. Всё это приводило к неоправданно высоким временным и трудовым затратам, связанным с отбором и редактированием загруженных в систему данных. Кроме того, между выходом издания в свет и поступлением библиографической информации о нём из РКП иногда проходило до полугода, что приводило к отказам и проблемам с поставкой.

Поэтому к 2018 г. назрела необходимость ввода библиографических записей в экспертную систему на основе данных, предоставляемых издательствами напрямую, а не после централизованной обработки. К сожалению, такая информация рассылается издательствами в основном в виде текстовых файлов или электронных таблиц Excel, то есть, является слабо структурированной и вариативной в зависимости от источника, а информационная система «Web-кабинет ученого» не обладает возможностями автоматизированного импорта информации из файлов подобного рода.

Отметим два аспекта этой задачи: с одной стороны, информационный поток изменился качественно, так как вместо одного источника стали использоваться несколько, причем весьма разнородных, хотя и более отвечающих тематической направленности библиотеки; с другой стороны, потребовалось технологически иное решение импорта внешней информации. В качестве решения была предложена процедура последовательной обработки данных на основе интеграции различных информационных систем: ввод данных в систему «ИРБИС», в которую встроены возможности импорта файлов из формата Excel; первичное аккумулирование, корректировка и селекция записей; экспорт данных в файл формата RUSMARC; импорт записей из файла RUSMARC в систему «Web-кабинет ученого»; окончательное редактирование набора записей и выставление их на экспертизу.

В ЦНБ УрО РАН программы семейства «ИРБИС» (Ассоциация ЭБНИТ, <http://www.elnit.org>) являются одними из наиболее ранних внедренных систем. В настоящий момент — это система «ИРБИС64», в составе ее серверной и клиентской части, и Web-сервер «Web-ИРБИС». Передача информации между системами «ИРБИС» и «Web-кабинет ученого» может осуществляться через файлы формата RUSMARC. Этот формат фактически является библиотечным стандартом обмена данными через файлы, но, в силу своей специфичности, не поддерживается офисными пакетами. Чтобы прочитать данные из файла такого формата требуются специализированные программы.

В системе «ИРБИС64» для обработки данных используются модули АРМ Администратор, АРМ Комплектатор и АРМ Каталогизатор. Средствами Администратора прайсы издательств из формата Excel загружаются в систему, при этом каждому столбцу исходной таблицы присваивается произвольная метка поля. Затем обработанный файл преобразуется в формат ISO2709, а созданный ISO-файл импортируется в базу данных системы (БД). Далее запускается АРМ Комплектатор, в котором с помощью режима «Переформатировать БД по заданному

соответствию меток» произвольным меткам полей присваиваются метки полей системы «ИРБИС». После завершения преобразования БД готова для дальнейшей работы: редактирования, вывода на печать, выгрузки в формат RUSMARC. Процесс преобразования данных в формат RUSMARC в системе «ИРБИС» является достаточно трудоемким и многостадийным. Однако в новой версии, «ИРБИС 2017», разработчиками предложено более быстрое решение данной задачи: загрузка прайсов непосредственно в АРМ Комплектатор, что существенно упрощает работу.

Предложенная технология оказалась вполне работоспособной. Приведем выборочную статистику для сравнения показателей Экспертной системы комплектования до и после изменений в алгоритме формирования наборов записей. В 2018 г. в экспертную систему комплектования произведено 6 загрузок (19 402 записи), загрузки № 1-2 прошли до модернизации системы, № 3-6 — после модернизации. На экспертизу представлено 1 439 записей. При изменении источника поступления библиографической информации для первичной обработки сократился объем обрабатываемых записей (приблизительно в 3-4 раза), поскольку появилась возможность в большинстве случаев сразу отсекаать издания, не соответствующие научному профилю и тематике исследований УрО РАН (рис. 1).

Такое сокращение, в целом, не привело к уменьшению общего объема выставляемых на экспертизу изданий и количества положительных оценок экспертов (рис. 2). Исключением может выглядеть загрузка №6, но, по нашему мнению, она вполне объяснима типичными сезонными изменениями количества изданий научной направленности. Кроме того, если такая тенденция окажется систематической, ее можно купировать расширением или корректировкой списка внешних источников записей.

В процентном отношении доля изданий, выставленных на экспертизу, от общего количества обработанных записей закономерно увеличилась и вместо 3-4% составила 11-23%. Процент положительных оценок экспертов на предложенные для экспертизы издания практически не изменился и варьируется в пределах 40-54% (табл. 1).

Таким образом, интеграция различных информационных систем при модернизации работы Экспертной системы комплектования ЦНБ УрО РАН позволила оптимизировать технологический процесс работы с системой на этапах предобработки библиографической информации и комплектования библиотеки рекомендованной книжной продукцией за счет сокращения временных и трудовых ресурсов на отбор научной литературы и селекции её постав-

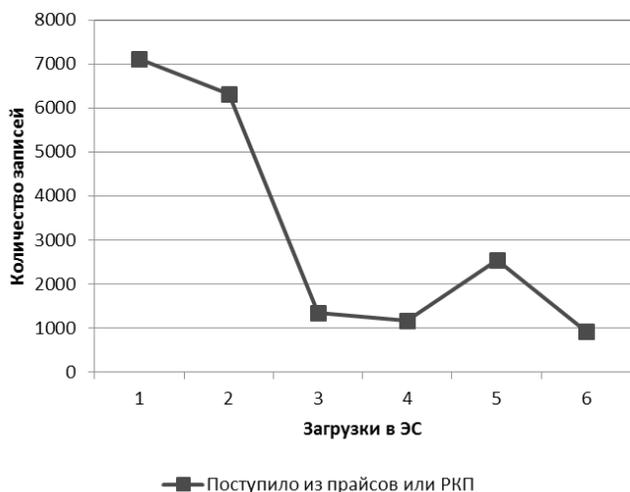


Рис. 1. Загрузки записей в Экспертную систему комплектования для первичной обработки в 2018 г.

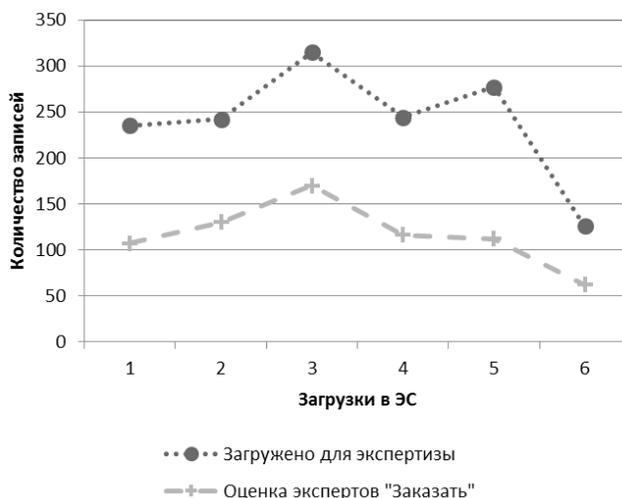


Рис. 2. Объемы загрузок записей в Экспертную систему комплектования и количество положительных оценок экспертов в 2018 г.

Таблица 1

Соотношение количества записей для обработки до и после модернизации экспертной системы комплектования

Период	Количество записей				
	Поступило из прайсов или РКП	Загружено для экспертизы	% загрузки для экспертизы от общей загрузки	Оценка экспертов «Заказать»	% положительных оценок от выставленных на экспертизу
2018 г.	19 402	1 439	7	697	48
Загрузка 1	7 119	235	3	107	46
Загрузка 2	6 305	242	4	130	54
Загрузка 3	1 348	315	23	170	54
Загрузка 4	1 156	244	21	116	48
Загрузка 5	2 544	277	11	112	40
Загрузка 6	930	126	14	62	49
Суммарно					
До модернизации системы (загрузки № 1-2)	13 424	477	4	237	50
После модернизации системы (загрузки № 3-4)	5 978	962	16	460	48

щиков. При этом изменение внешних источников информации при работе с Экспертной системой комплектования не привело к ухудшению качества проводимой экспертной оценки. Изменение технологии загрузки данных показало работоспособность такого решения, хотя оно и является, все же, достаточно паллиативным, а, следовательно, нужно искать другие возможные варианты межсистемного взаи-

модействия и импорта сторонних файлов, облегчающие работу с информацией библиотечной направленности. Отметим также, что в рамках интегрированного комплекса систем решена задача, которую было бы невозможно решить, используя какую-либо из информационных систем изолированно, то есть, проявился синергетический эффект.

Литература

1. Горбич Л.Г. Опыт внедрения экспертной системы комплектования в ЦНБ УрО РАН / Л.Г. Горбич, А.А. Живодеров // Новые технологии в информационно-библиотечном обеспечении научных исследований. Екатеринбург. — 2010. — С. 87—92.
2. Горбич Л.Г. Преимущества интерактивной электронной библиотеки (на примере информационной системы «Web-кабинет ученого») / Л.Г. Горбич // V Информационная школа молодого ученого: сб. науч. тр. / ЦНБ УрО РАН. Екатеринбург. — 2015. — С. 34—38.
3. Кирсанова А.И. Отражение информационных потребностей научных сотрудников Уральского отделения РАН в экспертных оценках интерактивной системы комплектования / А. И. Кирсанова // Информационное обеспечение науки: новые технологии. Екатеринбург. — 2016. — С. 42—52.
4. Костарева В. Е. Применение информационных технологий в текущем комплектовании ЦНБ УрО РАН на основе системы экспертных оценок / В. Е. Костарева // Информационное обеспечение науки: новые технологии. Екатеринбург. — 2012. — С. 147—152.
5. Оганова О.А. Проектная деятельность академической библиотеки в информационном обеспечении науки / О.А. Оганова, П.П. Трескова // Информационное обеспечение науки: новые технологии. Екатеринбург. — 2016. — С. 5—22.

REFERENCE

1. Gorbich L.G. Opyt vnedreniya ehkspertnoj sistemy komplektovaniya v CNB UrO RAN / L.G. Gorbich, A.A. Zhivoderov // Novye tekhnologii v informacionno-bibliotechnom obespechenii nauchnyh issledovanij. Ekaterinburg. — 2010. — S. 87—92.
2. Gorbich L.G. Preimushchestva interaktivnoj ehlektronnoj biblioteki (na primere informacionnoj sistemy «Web-kabinet uchenogo») / L. G. Gorbich // V Informacionnaya shkola mladogo uchenogo: sb. nauch. tr. / CNB UrO RAN. Ekaterinburg. — 2015. — S. 34—38.
3. Kirsanova A.I. Otrazhenie informacionnyh potrebnostej nauchnyh sotrudnikov Ural'skogo otdeleniya RAN v ehkspertnyh ocenках interaktivnoj sistemy komplektovaniya / A.I. Kirsanova // Informacionnoe obespechenie nauki: novye tekhnologii. Ekaterinburg. — 2016. — S. 42—52.
4. Kostareva V.E. Primenenie informacionnyh tekhnologij v tekushchem komplektovanii CNB UrO RAN na osnove sistemy ehkspertnyh ocenok / V.E. Kostareva // Informacionnoe obespechenie nauki: novye tekhnologii. Ekaterinburg. — 2012. — S. 147—152.
5. Oganova O.A. Proektnaya deyatel'nost' akademicheskoy biblioteki v informacionnom obespechenii nauki / O.A. Oganova, P.P. Treskova // Informacionnoe obespechenie nauki: novye tekhnologii. Ekaterinburg. — 2016. — S. 5—22.

