

# ОРГАНИЗАЦИЯ ИНФОРМАЦИОННОЙ РАБОТЫ

УДК 001.4:(0,0...)/(09).004.4

В. П. Леонов, Ю. П. Нюкша, В. П. Захаров

## Терминологические проблемы создания информационных комплексов по консервации документов

*Рассматриваются основные направления работы по решению задачи информационного обеспечения проблем сохранения и консервации документов в библиотеках: подготовки библиографических указателей, создания различных баз данных, разработки программного и лингвистического обеспечения.*

### 1. ОБЩИЕ ПРЕДПОСЫЛКИ

Библиотекой Российской академии наук (БАН) проводится новаторская, беспрецедентная по замыслу и размаху информационно-накопительная работа, касающаяся вопросов обеспечения сохранности документов в библиотеках. Она включает в себя подготовку библиографических указателей, создание различных баз данных (библиографических, фактографических, авторитетных), разработку программного и лингвистического обеспечения.

В основу работы положена процессная парадигма в библиотековедении, когда главным объектом библиотековедения рассматривается библиотечный процесс в системе документальных научных коммуникаций [1]. Подход к библиотековедению как к неотъемлемой системе социально-коммуникационных наук, служит отправной точкой в области прикладных исследований в книжном, архивном и музейном деле. Под таким углом зрения комплекс задач по обеспечению сохранности (сохранению) документов в библиотеках логично вписывается в систему библиотековедения в целом.

Задача информационного обеспечения проблем консервации имеет несколько составляющих и соответственно подразделяется по видам работ и стадиям разработки.

Под информационным обеспечением здесь понимается комплекс средств и методов, определяющих обработку и предоставление пользователям научной и технологической информации, а также сами информационные массивы. Соответственно, различные типы баз данных — библиографические, полнотекстовые, фактографические, материаловедческие — требуют своего инструментария. Точно так же различные этапы выполнения работы диктуют свои подходы и методы.

По разным причинам работа была по-своему новой для всех участников, и не все детали ее были уточнены с самого начала. Цели создания информационного комплекса, его адресаты, назначение, а соответственно, границы и структура формировались в процессе работы. Этой неопределен-

ности способствовало также отсутствие полноценной терминологической базы, обеспечивающей однозначное толкование процессов и явлений. Неустойчивое состояние терминологии особенно отчетливо проявляется при накоплении больших массивов однотемной, разноязычной информации. Более того, комплексная работа при общем возрастании требований к качеству информации обнаруживает расширение расхождений между профессионально-тематическими и даже национальными библиографиями, с одной стороны, и базами автоматизированных информационных систем — с другой.

Первый этап накопления материала сменяется его осмыслением и упорядочением. Решающую роль приобретает понятийная, содержательная сторона информационных процессов, которая под силу только квалифицированным специалистам, всесторонне образованным консерваторам. Однако именно в этой области заключены многочисленные сложности. Создаются ситуации, для которых приходится находить временные компромиссные решения, принимая во внимание состояние консервации как научной дисциплины. Для профессионалов не всегда очевидны и применимы однозначные решения. В рамках самой консервации, особенно если рассматривать ее через призму библиотековедения, существует еще немало проблем и нерешенных вопросов.

### 2. ПЕРСПЕКТИВНЫЕ НАПРАВЛЕНИЯ

Постепенно в ходе работы по созданию в БАН информационной системы определились ее составляющие. Это библиографическая база данных с массивом типовых запросов, полнотекстовая документальная информационная система, фактографический банк данных по сохранению документов, база данных по микроклимату в хранилищах. Названные системы имеют свое информационное и лингвистическое обеспечение. В то же время поставленные задачи тесно связаны друг с другом — в том числе и через лингвистическую компоненту. Из осознания особой актуальности и, можно ска-

зать, изначальности терминологических проблем и появилась эта статья.

На данный момент наибольшее развитие получила библиографическая подсистема, "побочным продуктом" которой является библиографический указатель. Большие объемы и универсальное целевое назначение библиографического материала приводят к необходимости выделения наиболее перспективных и полезных аспектов деятельности. Одним из таких направлений является реферирование и перевод на иностранные языки отечественной литературы по сохранению документов. В оригинале она практически недоступна для зарубежных специалистов. Подобное неудовлетворительное положение умаляет в международных кругах представления об уровне работы российских специалистов, формах и направлениях их деятельности. В итоге сужаются возможности эквивалентного обмена и распространения литературы по консервации документов.

Неоспоримо, однако, что эффективная работа возможна только в случае сопровождения библиографических описаний поисковыми аннотациями, составленными современным профессиональным языком с использованием специфических значений слов корпоративной лексики. Перевод требуется отнюдь не дословный, а смысловой. При этом качество перевода становится в зависимости от применения всего терминологического спектра данной отрасли знания и правильной ориентации в частных случаях его употребления.

В конечном счете, полезность и перспективность данных по консервации документов на международном уровне состоит в умелом предоставлении зарубежным пользователям русскоязычной специальной литературы. Кроме того, одного реферирования может оказаться недостаточно. Пользу принесет также подготовка авторитетных данных, включая разнообразные сведения, расширяющие информацию об авторе, о предмете исследования, методических и социально-психологических подходах к проблеме. Безусловно, что такое раскрытие тематики позволяет более предметно отвечать на вопрос пользователя в гибкой формулировке и освещать его с большей степенью глубины.

### 3. ТЕРМИНОЛОГИЧЕСКОЕ ПРОСТРАНСТВО

Известно, что слова или сочетания слов специфичны для каждой отрасли знания и не только влияют на внешнюю оформительскую сторону информационных комплексов, но и во многом определяют их содержательные пределы и допуски. Классификационная система организации информационного блока сведений и сопровождающий его тезаурус непосредственно связаны между собой и смыкаются на терминологическом поле не только консервации, но и культуры в целом. Более того, современная постановка вопросов сохранения культурных ценностей выходит на терминологическое поле естественных фундаментальных наук — физики, химии, биологии, а также экологии, химии полимеров, различных производственных технологий — полиграфии, бумагоделательного производства, наконец, сферы санитарно-профилактической.

Классификационные схемы консервации всегда отражают сложившиеся исторически эмпириче-

ские решения. Именно поэтому они нуждаются в теоретической проработке для определения терминологически выверенной структуры базы данных. При этом не менее значима семантика всех ее элементов с указанием тождеств и различий в определениях и всей лексике описаний. Терминологический анализ библиографических, фактографических и аналитических описаний постепенно выявляет синонимические элементы, несущие равную смысловую нагрузку, но представленные разными лексическими формами.

При создании информационных систем большого объема должно быть определено постоянство форм. В первую очередь, это форматы, но проблемы изменчивости лексики также должны быть учтены. Справедливость такого утверждения отчетливо проявилась при составлении электронных каталогов, которыми наши зарубежные коллеги овладели ранее и полностью прочувствовали вредность частых библиографических и лексических изменений. В то же время при совершенствовании и унификации отдельных элементов надлежит избегать неоправданного разнообразия, которое в большой степени отражается на тезаурусе. При составлении четкой и экономичной модели терминологический анализ проводится с учетом разноязычия. Известны трудности совмещения профессиональной русской терминологии с соответствующими эквивалентами на других языках и создаются трудности при их объединении в одном ряду [2]. В консервации такой феномен проявляется со всей очевидностью, тем более, что эта область знаний находится на стыке гуманитарных и естественных наук. Именно в этом случае пренебрежение терминологией может завести в джунгли полной хаотичности. Вполне вероятна ситуация, при которой люди одной специальности перестанут понимать друг друга.

Наиболее эффективным механизмом упорядочения лексики является стандартизация. Она наводит мосты между библиотечно-библиографическими дисциплинами и другими отраслями знания. Типичным примером является термин "документ". По ГОСТ 16487-83 [3] он определен как "материальный объект с информацией, закрепленной созданным человеком способом для ее передачи во времени и пространстве". В этой формуле содержится подтверждение научного дуализма самой сути консервации, которая имеет дело с двусторонней единицей, не отражающей ее собственную сущность и именуемой носителем (любой) информации. Заметим, что термин "документ" после тщательной проработки терминологии на основе базы фонда стандартов, закреплен за всеми изданиями и рукописями, хранящимися в библиотеках, и употребляется при стандартизации библиотечно-библиографических процессов. Однако в библиотеках, не только в языке общения, но и в специальной литературе, он воспринимается неохотно или вовсе заменяется другими терминами.

Существует ГОСТ 7.48-90 [4] (в настоящее время пересматривается в связи с истечением срока действия), в котором установлены применяемые в науке, технике, производстве термины и определения понятий в области консервации документов. Всего стандартизовано 67 терминов, определения которых удалось согласовать со специалистами страны. В обиходе находится значительно

больше число терминов, и их арсенал постоянно увеличивается по мере развития технологии консервации, а также переосмысления ее теоретических основ.

Таким образом, терминологическая база консервации находится в развитии и модернизации. Именно поэтому при соблюдении определенной корректности ситуацию надлежит отслеживать, не допуская дезорганизующего воздействия устаревших понятий на структуру информационной системы. Иными словами, терминологические процессы, находясь в естественном движении, во многих случаях в каждый данный момент приводят к несходным определениям.

#### 4. ОСОБЕННОСТИ ТЕРМИНОЛОГИИ В ОБЛАСТИ КОНСЕРВАЦИИ

Анализ имеющихся формулировок показывает, что в области консервации превалирует не нормативная, а аспектная терминология, присущая отдельным подразделениям консервации, целым учреждениям и даже регионам. К сожалению, достичь однозначного толкования терминов помимо стандартизации не удается в силу некой косности, недостаточной осведомленности, неправомерной оценки собственных представлений некоторыми специалистами, а в целом благодаря пренебрежению аналитическим подходом. В итоге терминологическое поле консервации нередко оказывается сбивчивым.

Скопившиеся неточности поступают в обращение, и их негативное влияние особенно усиливается при описании одних и тех же явлений на различных языках. Поэтому в обиходе существует разнородная смесь толкований, синонимов и омонимов. Процесс составления классификационных схем и распределения публикаций по разделам при таких обстоятельствах затрудняется. Особую сложность представляет классификация публикаций на основе одних заголовков, названий статей или книг.

Например, термин *консервация*, *conservation* стал настолько многоаспектным и собирательным, что по ГОСТ 7.48-90 его пришлось разделить на три составляющие: “режим хранения”, “стабилизация” и “реставрация”. В них укладываются большинство известных технологических процессов. Однако в зарубежных источниках, не придерживающихся российского стандарта, термин “*conservation*” используется произвольно согласно представлениям авторов. В тех случаях, когда проступает явная несовместимость в одном слове слишком разных технологических процессов, вводят как бы разъясняющие дополнения, такие, например, как *conservation treatment*, что вполне могло бы быть заменено термином *stabilization*. Однако и этот термин на различных языках используется в отношении непохожих технологий. В русских источниках (иногда также и в английских) под стабилизацией понимается прекращение активно протекающих химических и биологических процессов, приводящих к повреждению, разрушению и даже полной деструкции материальной основы документа. Иное употребление термина *stabilization* у немецких авторов, которые его приравнивают толь-

ко к одной из операций реставрации — приданию формы документу, в том числе и путем дополнения недостающих частей, т. е. оставляя за стабилизацией функцию механического восстановления формы листа.

Существуют также условности в употреблении обычных слов. Например, состояние воздушной среды — *air condition*, а состояние бумаги — *paper state*, точно так же как состояние других материалов книги. Суперобложка может обозначаться как *dust cover*, *wrappier*, *jacket* и др. А обычное слово *case*, встречающееся чаще всего как “случай”, обозначает в библиотечной практике и *ящик*, и *чехол*, и *футляр*, и *коробка*, и *витрина*, наряду с тем, что для этих слов имеются и другие специфические обозначения, например, *drawer*, *box*, *chest*, *tray*, *cover* и даже *sleeve*. В связи с этим трудно переоценить значение многоязычных сравнительных таблиц. Дело их составления неблагоприятное, но избежать его не удастся и практически предстоит выполнить заново: все существующие ныне терминологические словари по консервации таят в себе коварные погрешности, приводящие к казусам и парадоксам. Так, при ошибках в толковании слов поиск по тезаурусам ведет к выпадению целых “кустов” — больших объемов литературы, отвечающей на поставленный вопрос. При этом несомненно важна позиция и осведомленность ведущего поиска: пользуется ли он только своими собственными представлениями или применяет другие, хотя бы наиболее известные концептуальные схемы.

Углубление и расширение работ по созданию информационных систем, охватывающих вопросы консервации, приводит к необходимости все основательнее вникать в проблемы терминов и определений. В связи с этим неизбежны также трудности преодоления психологических барьеров.

#### 5. ИЕРАРХИИ ПОНЯТИЙ

Особое место занимает вопрос о соподчиненности понятий, существующих в области консервации документов. Это касается и сути процессов, протекающих в документах, в хранилищах, и закономерностей функционирования системы хранения и обеспечения доступа к фондам библиотеки. Все реально протекающие в жизни документа процессы находят отражение в информационных системах. Для правильной организации тезаурусов и классификационных схем устанавливается родовидовая подчиненность терминов, уровень понятий, их ступенчатость, взаимосвязь, взаимозависимость. Так, например, термин *консервация* по ГОСТ 7.48-90 рассматривается как родовой по отношению к терминам *режим хранения*, *стабилизация* и *реставрация*, а термин *обеспечение сохранности документов* охватывает все виды работ, проводимых в библиотеке для поддержания фонда в доступном пользователю состоянии. В настоящее время этот термин часто заменяется на *сохранение*.

Весь тезаурус выстраивается в иерархической системе понятийных уровней с выявлением синонимии действительной и мнимой. Мнимая синонимия, к сожалению, встречается не столь редко, так как возникает на основе разного толкования одних и

тех же терминов и определений различными людьми. Чтобы избежать разночтений в отсутствие общепризнанной системы иерархии, важно последовательно придерживаться одного порядка, ибо у пользователя не должно возникнуть смещения понятий.

В усилении четкости нуждается сущностное определение слов в тезаурусе. При поиске сущностно близкие слова должны быть легко выявляемы и сопоставляемы. Например, такие понятия, как *налет плесени, обрастание грибами, грибные колонии, колонизация, микробиологическое, а также микологическое повреждение* по существу явления одного порядка, но без специальных отсылок будут так разобщены, что эффективность пользования тезаурусом значительно снизится.

Аналогичная ситуация и в области обозначения конструктивных элементов книги. В полиграфических стандартах и словарях узаконено применение понятий: *приклеенные элементы, крытье в роспуск, бесшовное скрепление, распашная иллюстрация* и др. В библиотеках нередко такие определения заменяют на более благозвучные с точки зрения русского языка. Однако такой подход ни в каком отношении пользы не приносит, наоборот, вносит путаницу в тезаурус, затрудняя поиск однотемных работ, опубликованных в библиотечных и в полиграфических изданиях.

## 6. НАПРАВЛЕНИЯ ДЕЙСТВИЙ

К настоящему времени накоплен огромный информационный задел по консервации и по сохранению документов в целом. Такие работы ведутся в большей или меньшей степени во всех крупных центрах по консервации документов. Параллельно прилагаются усилия по выработке концептуальной программы информационной системы. Успех дела во многом определяется в ее содержательной сфере, связанной в большой мере с четкой отработкой терминов и определений.

Впереди несомненное расширение объема и частичный пересмотр терминологических возможностей, на которые было бы позволительно опереться как на аргументированный и глубоко проработанный материал. Принимая во внимание объем информационной системы, ее сущностную неоднородность и разноязычие, ожидать скорого и окончательного решения проблемы не приходится. В настоящий момент единственно доступным является поэтапное ведение работы.

Анализ ситуации показывает, что существующие аспектные терминологические словники служат соединительными, переходными к нормативной документации. Поэтому любое принимаемое решение является временным по мере дефиниции, сущностного толкования слов, содержащихся в информационной системе. Предстоит выработать механизм постепенного перехода от произвольных аспектных к научным классификационным системам на основе выверенных моделей терминов и определений. Формирование информационных систем протекает под влиянием многих факторов,

связанных непосредственно с самой работой и сопутствующих ей. В связи с этим следует обратить внимание на следующие обстоятельства.

\* Пока еще рано переоценивать информацию по консервации, добываемую посредством системы Интернет. В нее попадает не все и часто не самое главное. Так что надобность в создании собственных систем остается и сохранится еще на продолжительное время.

\* Печатные реферативные издания по консервации либо не поступают в наши библиотеки, либо отражают мало доступную, преимущественно ведомственную литературу.

Особого внимания требует вопрос сочетания библиографических данных и фактографического материала на различных языках, в том числе на русском. Эту проблему представляется возможным решить следующим образом. Все русскоязычные работы, как уже отмечалось, надлежит снабжать переводом названий и поисковыми аннотациями на одном или даже нескольких иностранных языках. Все иноязычные подборки по темам надлежит дополнять соответствующими по содержанию русскоязычными при условии обязательного понятийного аннотирования с указанием разделов, новизны, оригинальности и основных выводов.

В информационной системе полезно отвести место библиографии второй степени, которую не сложно сформировать из внутрикнижных и внутривстречных библиографий. Принимая во внимание существование не устоявшейся ситуации, разумнее всего в настоящее время установить этапность в создании информационной системы по консервации.

В БАН в настоящее время находится в стадии завершающей разработки информационно-поисковый тезаурус (ИПТ) по сохранению библиотечных фондов, предназначенный для индексирования документов и запросов в базах данных (БД).

Основа словника ИПТ сформирована из ключевых слов, представленных в поисковых образах документов, уже введенных в базу. Из этого массива около 1,5 тыс. документов было индексируемо вручную, оставшиеся — автоматически: из текстов заглавий и аннотаций выделялись ключевые слова и при минимальном отсеке неинформативных слов (предлоги, союзы, артикли и некоторые др.) программными средствами формировались поисковые образы документов. Для нормализации лексических единиц (ЛЕ), установления синонимических и иерархических отношений в тезаурусе использованы отечественные стандарты. Использовались также тезаурусы, словари, учебная и специальная литература, некоторые рубрикаторы информации по тематике данной предметной области. Собранный лексический материал оценивали, верифицировали и в ряде случаев существенным образом пополняли путем экспертной оценки специалистов БАН в области консервации документов.

Тезаурус содержит три части:

(1) алфавитный лексико-семантический указатель ключевых слов и дескрипторов (см. фрагмент):

**COMBAT MESUARES**

- s. Decontamination  
Деконтаминация  
Истребительные меры борьбы  
с биологическим фактором  
Обеззараживание
- n. DERATISATION  
DISINFECTION  
DESINSECTION  
ELECTROMAGNETIC TREATMENT  
FUMIGATION  
HIGHFREQUENCE TREATMENT  
INERT GASES TREATMENT  
NONTOXIC TREATMENT  
X-RAY TREATMENT  
Дезинфекция  
Высокочастотная обработка  
Обеззараживание  
Обработка инертными газами  
Фумигация  
Электромагнитная обработка
- b. HOUSEKEEPING  
Гигиеническая обработка

**FREEZE DRYING**

- s. Blast freezers  
Freezedrying  
Freeze-drying  
Freeze dryers  
Freezing  
Mechanical freezers  
Redesigned freezers  
Замораживание  
Морозильные камеры  
Сублимационная сушка  
Сушка вымораживанием
- b. DRYING  
Сушка
- a. DEFROSTING  
Размораживание

**FUMIGATION**

- s. Desinfection  
Дезинфекция  
Фумигация
- b. COMBAT MESUARES  
Истребительные меры борьбы  
с биологическим фактором

(2) списки идентификаторов — терминов, выражающих единичные понятия. В эту часть ИПТ, вынесенную для удобства работы за пределы основного лексического массива тезауруса, входят наименования организаций, отдельных наиболее ценных памятников культуры и письменности, временные и географические привязки объектов;

(3) иерархический указатель.

В написании ключевых слов и дескрипторов используются буквы русского и латинского алфавитов, без диакритики, подстрочных, надстрочных и т. п. знаков. Словарные статьи упорядочены по алфавиту.

Тезаурус подготовлен в соответствии с ГОСТ 7.25-80 [5] с учетом некоторых рекомендаций из ГОСТ 7.24-90 [6].

Основными особенностями ИПТ по сохранению документов является выбор английского языка в качестве основного в тезаурусе, а также широкая интерпретация отношения синонимии (эквивалентности терминов). Выбор английского языка как языка индексирования и основного языка при формировании словарных статей был сделан по прагматическим соображениям. На первом этапе создания БД по сохранению документов обрабатывались преимущественно английские тексты и индексация английскими ключевыми словами во многом решала иногда непростую проблему их перевода на русский язык. "Русификация" ИПТ и перевод на русский язык терминов в поисковых образах документов является непростой, но чисто технической проблемой, решаемой программными средствами, разработка которых ведется в БАН.

Итак, подведем предварительные итоги.

Первоначально тезаурус и классификационные схемы строятся по временной терминологической модели на основе ключевых слов в заголовках работ. Накопление материала можно считать пройденным этапом, но в принципе эта работа бесконечна. Далее неизбежен этап глубокого сущностного анализа лексики с учетом семантики, струк-

туры и функций слов и понятий. Здесь уже речь идет о привлечении других авторитетных данных и фактографического материала.

При корректировке словника, рубрикатора, классификационной схемы используются возможности самой компьютерной системы с дополнительным введением в оборот неучтенных, но сущностно близких слов, а также выявление расхождений в их понятиях.

Непременное условие совершенствования системы — постоянный терминологический контроль, являющийся одновременно и способом умножения ее достоинств.

Основной целью и главной оценкой качества информационной системы по консервации является удачное объединение разнотемной, разножанровой и разноязычной литературы, нужной и интересной всем специалистам библиотек, музеев и архивов.

**СПИСОК ЛИТЕРАТУРЫ**

1. Леонов В. П. Библиотечно-библиографические процессы в системе научных коммуникаций. — СПб: Б-ка Рос. Акад. наук, 1995. — 140 с.
2. Бахтурина Т. А., Сукиасян Э. Р. Что же произошло с терминологическими стандартами? // Научные и технические библиотеки. — 1997. — № 9. — С. 3-12.
3. ГОСТ 16487-83. Делопроизводство и архивное дело. Термины и определения. Введ. 01.01.85.
4. ГОСТ 7.49-90. СИБИБД. Консервация документов. Основные термины и определения. Введ. 01.01.91.
5. ГОСТ 7.25-80 (СТ СЭВ 174-75). СИБИБД. Тезаурус информационно-поисковый одноязычный. Правила разработки, структура, состав и форма представления. Введ. 01.01.82. — Изм. 1. Введ. 01.07.87.
6. ГОСТ 7.25-80 (СТ СЭВ 6574-89). СИБИБД. Тезаурус информационно-поисковый многоязычный. Состав, структура и основные требования к построению. — Введ. 01.01.91.

*Материал поступил в редакцию 02.11.99.*