

УДК 002.6:54

А.А.РАГОЙША

ПАТЕНТНЫЕ РЕСУРСЫ WORLD WIDE WEB

Review of free patent databases available in the Internet. A brief description (addresses, subject coverage, search engines) of the resources provided by USPTO, Questel-Orbit, IBM Patent Server, MicroPatent, SPO, etc. is given. The paper contains a manual on searching for information in QPAT-US database.

Обязательным этапом любого прикладного исследования является патентный поиск. По некоторым оценкам, патенты содержат 85–90% производимой в мире технологической информации, причем лишь 5–10% этих сведений позже публикуется в иной научной литературе [1]. Поскольку общее число патентов, изданных двумя сотнями национальных и международных патентных бюро, уже достигло 50 млн, в мировой практике в настоящее время широко

применяется компьютеризация поисковой работы. Исходный материал в форме электронных баз данных поступает к читателю на компакт-дисках либо предоставляется через систему телекоммуникаций коммерческими онлайн-службами *STN*, *Questel-Orbit*, *DIALOG*, *Data Star* и др. [2–4]. Пользование такими услугами требует больших финансовых затрат, поэтому нами были изучены бесплатные патентные ресурсы *World Wide Web*. С учетом характерной для Интернет проблемы достоверности распространяемой информации в настоящем обзоре мы рассмотрим только те базы данных, которые поддерживаются официальными патентными организациями и фирмами, имеющими устойчивую репутацию в научной среде.

По степени информативности и систематичности материала эти ресурсы можно разделить на три категории:

1. Патенты США (практически исчерпывающие сведения за последние 25 лет).
2. Патенты *European Patent Office* (библиография и рефераты за последние 20 лет).
3. Остальные патенты (фрагментарные сведения с явно выраженной тенденцией к увеличению их количества).

Несколько организаций в разном объеме открывают свободный доступ к патентам США. Поисковые программы соответствующих баз данных подобны, но не идентичны, и при выборе конкретного источника информации читателю следует учитывать загруженность сервера, свой опыт работы и цели конкретного библиографического исследования.

US Patent and Trademark Office (USPTO)

Center for Networked Information Discovery and Retrieval (CNIDR)

<http://patents.uspto.gov/>

<http://patents.cnidr.org/>

<http://app.cnidr.org/>

База данных *US Patent Bibliographic Database* состоит из титульных страниц (библиографических описаний, рефератов и списков литературных источников) патентов США, изданных после 01.01.76. Материал титульной страницы распределен по 27 полям, что представляет ценность при планировании узконаправленного информационного поиска. Отдельные документы объединены системой гиперсвязей по критерию взаимного цитирования. Перечисленные выше серверы характеризуются приемлемой скоростью работы и могут быть рекомендованы для проведения экспресс-анализа американской патентной литературы.

Поисковое задание записывается на одном из трех предлагаемых бланков: *Patent Number Search* - при извлечении патентов по их номерам.

Boolean Search – при поиске информации по признаку наличия термина (комбинации терминов) в одном (двух) поле (полях) либо в титульной странице в целом; типы полей выбираются в меню.

Advanced Search – при информационном поиске по признаку наличия терминов в любом наборе полей: в этом случае условные обозначения полей записываются в самой поисковой фразе (тип не указывается, если поиск термина проводится по всей титульной странице).

Поисковая фраза формулируется с использованием логических операторов *AND*, *OR*, *ANDNOT*, а также *XOR*. Допускается сокращение правой части слова (знак *) и поиск строки символов (во фразе, обозначаемой двойными кавычками).

Пример 1. Найти сведения о покрытиях, запатентованных автором *Taylor*. Заполняем бланк *Boolean Search*:

- Указываем временной интервал: в поле *Select Database* отмечаем *All*.
- В поле *Term 1* записываем *coating**, в меню *Field 1* выбираем *Any Field*.
- В меню *Operator* выбираем *AND*.
- В поле *Term 2* записываем *Taylor*, в меню *Field 2* выбираем *Investor Name*.

Пример 2. Найти патенты *General Electric Company* на силикатные покрытия, не содержащие фосфаты.

Заполняем бланк *Advanced Search*:

■ Указываем временной интервал: в поле *Select Database* отмечаем *All*.

■ В поле *Query* записываем поисковую фразу:

and "General Electric" and coating and silicate* andnot phosphate**.

Список извлеченных патентов по желанию пользователя компоуется либо в хронологическом порядке, либо по степени соответствия поисковому заданию (*relevance ranking*).

Отметим, что с Главной страницы сервера *USPTO* есть переход к базе данных *AIDS Patent Database*, содержащей полные тексты патентов США, Японии и европейских стран по тематике СПИД.

Questel-Orbit

<http://qpat2.qpat.com/>

После предварительной онлайн-регистрации читателю разрешается работать с частью базы данных *QPAT-US*, а именно с титульными страницами патентов США, изданных после 01.01.74 г. (гиперсвязь *Abstracts Database*). Среди прочих этот информационный источник выделяется своей удобной и совершенной поисковой программой, поэтому структуру *QPAT-US* рассмотрим более подробно.

1. Основная поисковая страница (*main search page*).

Меню *Default Field* предназначено для указания поля, в котором предполагается вести поиск; пункт *Free Group* соответствует сумме всех перечисленных полей.

Временной интервал поиска (по дате опубликования патента) выбирается в меню *Date Range*.

Поисковое задание формулируется в редактируемом поле бланка. Предусмотрены два способа построения поисковой фразы: на естественном языке либо с использованием логических операторов *AND*, *OR*, *NOT*, *ADJ*, *Win*, *NEAR/n*. Группа слов без операторов по умолчанию объединяется оператором *OR*. Разрешен поиск строки символов (во фразе, обозначаемой одинарными кавычками). Применение шаблона (знак ? заменяет один символ) и функции сокращения (знак * заменяет любое количество символов) допускается в начале, середине и конце слова. Поиск ведется в режиме *stemming* (автоматическое дополнение задания вариантами грамматических форм); для отключения этого режима в конце термина записывается знак #.

Пример. Найти патенты на использование фосфорсодержащих солей алюминия и хрома как связующих.

В поисковой фразе учитываем особенности американского и британского написания термина "алюминий" и смысловую многовариантность термина "фосфаты" (гидро-, поли-, металло- и т.п.):

*((aluminum or aluminium or chromium) adj *phosphate*) and (binder* or adhesive*)*

Читатель может вызвать на экран (кнопка *Expand Terms*) вспомогательные указатели:

Fuzzy Match Adviser – список близких по написанию слов;

Statistically Related Adviser – перечень терминов, предположительно контекстно близких;

Database Dictionary Adviser – алфавитный список всех слов, встречающихся в базе данных.

Указатели, во-первых, помогают концептуально уточнить область поиска, во-вторых, информируют об опечатках, возникших при сканировании и последующем распознавании текстов исходных документов. В частности, *QPAT-US* содержит такие необычные формы термина "алюминий", как *aluminum*, *aluminium*, *aluminium*, *alumintum*.

Кнопка *Submit Query* включает процесс поиска информации; по его завершении в группе *Search Sets* появляется новая гиперсвязь, ведущая к списку извлеченных патентов.

2. Список патентов (страница *view self*).

Список извлеченных патентов компонуется по степени их соответствия поисковому заданию (*relevance ranking*). Гиперсвязи, начинающиеся от элементов списка, направлены к титульным страницам конкретных патентов.

Для предварительного ознакомления с материалом удобно руководствоваться следующим алгоритмом работы: (а) переключатель *Default View* поставить в положение *Abstract*; (б) нажать кнопку *Batch View* – в результате на экран выводится один файл, содержащий все рефераты.

3. Окно с титульной страницей патента.

Для извлечения патентов, тематически связанных с данным, выбирается один из пунктов меню, находящегося в верхней либо нижней части окна. Программа использует следующие критерии поиска:

SimilarAbs. В качестве поисковой фразы используется текст реферата. Поиск проходит медленно, причем читатель получает много совершенно посторонней информации. Интересно, что в результате этой операции извлекаются полные тексты патентов, в обычном режиме доступные только подписчикам.

Backward. Извлекаются патенты, на которые ссылается данный документ.

Forward. Извлекаются патенты, ссылающиеся на данный документ.

Network. Извлекаются патенты-предшественники (операция *Backward*), после чего к каждому из них применяется операция *Forward*.

NarrowCls. Выбор патентов того же патентного класса/подкласса, который объявлен первичным для данного патента.

BroadCls. Выбор патентов тех же патентных классов/подклассов, к которым отнесен данный патент.

IBM Patent Server

<http://www.patent.ibm.com/>

База данных *U.S. Patent Database* выделяется тем, что содержит полные материалы (тексты и иллюстрации) патентов США, изданных с 1974 г. (патенты 1971 г. представлены частично). Титульные страницы патентов и патентные формулы поступают к читателю в формате.html; выбрав гиперсвязь *View Images*, можно получить отсканированное изображение всех листов патента, хранящееся в графическом формате.

В связи с большой загруженностью сервера и медленной скоростью переноса графических файлов этот интереснейший информационный источник целесообразно использовать только на последних стадиях поиска – для извлечения полного комплекта материалов патента по уже известному его номеру.

Для записи задания база данных имеет несколько бланков: *Search* (поиск по слову или группе слов), *Patent Number Search* (поиск по номерам патентов), *Boolean Text Search* (поиск в одном, двух или всех текстовых полях; названия полей отмечаются в бланке) и *Advanced Text Search* (поиск в 1–6 текстовых полях, их названия отмечаются в бланке или же записываются в поисковой фразе). В двух последних бланках поисковая фраза формулируется с использованием операторов *AND*, *OR*, *NOT*, *NEAR/n* и др.; следует, однако, учитывать ее специфический синтаксис (пример: *ion <ORDER> <NEAR/1> exchanger*). Применение шаблона (знак ? заменяет один символ) и функции сокращения (знак * заменяет любое количество символов) допускается в начале и конце слова. Поиск по умолчанию ведется в режиме *stemming*; для отключения этого режима перед термином записывается оператор *<WORD>*. Список извлеченных патентов компонуется по степени их соответствия поисковому заданию.

MicroPatent

<http://www.micropat.com/>

После предварительной онлайн-регистрации читатель получает право работать с тремя базами данных: *US Patent Search/Online Gazette*; *World Patent Search* и *This Week's FullText*.

US Patent Search/Online Gazette содержит титульные страницы, первые пункты патентных формул (*Claim*) и по одной иллюстрации из патентов США, опубликованных в официальном еженедельнике *USPTO Gazette* в течение последнего года.

World Patent Search является единственной бесплатной базой данных в WWW, предоставляющей систематическую информацию о неамериканских патентах. Здесь можно ознакомиться с титульными страницами изданных в 1978–1998 гг. патентов *European Patent Office*, а также с титульными страницами патентов США, изданных в 1976–1998 гг. Во время подготовки статьи к печати база данных находилась на стадии бета-тестирования.

Для извлечения информации из указанных баз данных используется три типа бланков: *Patent Number Search* (поиск по номеру патента), *Simple Search* (поиск терминов в названии, реферате, списках авторов и владельцев) и *Fielded Search* (поиск по всем полям). Поисковая фраза для *Simple Search* может быть сформулирована с использованием операторов *AND* и *OR*; в *Fielded Search*, кроме того, могут применяться *ADJn*, *NEARn* (в поле *Title or Abstract*), *TO* (в числовых полях) и некоторые другие. Группа слов без операторов по умолчанию объединяется оператором *AND*. Разрешен поиск строки символов (во фразе, обозначаемой двойными кавычками). Применение шаблона (знак ? заменяет один символ) и функции сокращения (знак * заменяет любое количество символов) допускается в начале, середине и конце слова.

Пример. Найти патенты, опубликованные с 01.01.97 по 31.03.98, на покрытия, отвердевающие при температурах, близких к комнатным.

При составлении задания учитываем, что искомый текст может содержать такие фрагменты, как *low temperature curing*, *cured at very low temperature* и т. д.

Заполняем бланк *Fielded Search*:

■ в поле *Issue / Publication Date* записываем
19970101 to 19980331

■ в поле *Title or Abstract* записываем
coating and (cur* near3 ("low temperature" or "room temperature"))*

База данных *This Week's FullText Patents* содержит полные тексты патентов США, изданных в течение последней недели. Поисковый бланк этой базы данных имеет одно редактируемое поле для записи задания (использование логических операторов разрешено).

Пример записи задания:

refractory and (ceramic or cement)

Одной из задач, которую приходится решать в ходе информационного поиска, является извлечение патентов, объединенных общей тематикой, по их классификационному коду. В большинстве уже рассмотренных баз данных допускается использование номера класса/подкласса в качестве элемента поискового задания. В отдельных случаях к серии родственных патентов можно перейти по гиперсвязи, которая начинается от номера класса/подкласса, приведенного в извлеченном материале. Для независимого же нахождения классификационного кода США как исходного поискового параметра следует прибегать к услугам специализированных баз данных, перечисленных ниже.

EDS Shadow Patent Office (SPO)

http://www.spo.eds.com/spo/open_bin/gen_free.pl/man_class

Списки классов и подклассов (даны коды и названия).

http://www.spo.eds.com/spo/open_bin/gen_free.pl/display_class

В поисковом бланке записывается код класса (класса и подкласса); в результате поиска извлекается описание класса и подкласса.

http://www.spo.eds.com/spo/open_bin/gen_ospomail.pl/e081

Поисковое средство для нахождения списка классов/подклассов и списка патентов по признакам, указанным в поисковой фразе. Задание формулируется на естественном языке (до 1000 слов). Для уточнения задания можно воспользоваться дополнительными полями: *Keywords* (перечень слов, несущих особую смысловую нагрузку) и *Anti-Keywords* (перечень терминов, которые должны отсутствовать в извлекаемых записях). База данных содержит патенты (номер, название, дата) с 01.01.95.

http://www.spo.eds.com/spo/open_bin/gen_ospomail.pl/e091

Поиск тематически близких патентов. По номеру патента определяется его место в классификационной системе (классы и подклассы) и извлекается набор родственных патентов (номер, название). Для уточнения задания можно воспользоваться дополнительными полями: *Keywords* (перечень слов, несущих особую смысловую нагрузку) и *Anti-Keywords* (перечень терминов, которые должны отсутствовать в извлекаемых записях). База данных содержит патенты с 01.01.95.

USPTO, CNIDR

<http://patents.cnidr.org/CLASSES/classes.html>

(либо гиперсвязь *Browse US Patent Classes* с Главной страницы сервера USPTO)

Текст Руководства по системе классификации патентов США (*US Manual of Classification*). От списка классов гиперсвязи направлены к подробным описаниям классов и подклассов.

<http://aids.uspto.gov/AIDS/access/search-class.html>

(либо гиперсвязь *Search US Manual of Classification* с Главной страницы USPTO)

Поисковое средство для нахождения списка классов по признакам, указанным в поисковой фразе. Задание формулируется как набор терминов и логических операторов. Поиск не контекстный, а только по факту наличия терминов в соответствующей главе *US Manual of Classification*. Даются подробные описания классов и подклассов.

Как уже отмечалось, единственным крупным источником сведений о патентах, изданных за пределами США, является *World Patent Search* фирмы *MicroPatent*. В настоящее время в стадии становления находится ряд баз данных, которые уже теперь могут представлять интерес для читателя.

Derwent

<http://www.derwent.com/plweb-cgi/fastweb?searchform+view1>

(либо <http://www.derwent.com/resource/resource.html> и гиперсвязь *Patent Discovery*)

Библиографические сведения (без рефератов) о патентах, полученных от 40 патентных организаций и включенных в базы данных фирмы *Derwent* в течение последних трех недель. Следует учитывать, что здесь приводятся не оригинальные названия патентов, а переработанные и расширенные экспертами *Derwent* (что в некоторой степени компенсирует отсутствие рефератов).

Japanese Patent Office

<http://www2.jpo-miti.go.jp/default.htm>

Библиографические сведения (с рефератами) о патентах Японии, изданных за период с 01.11.96 по 30.06.97.

Canadian Intellectual Property Office

<http://strategis.ic.gc.ca/cgi-bin/patent/searche.pl>

Титульные страницы (без рефератов) патентов Канады, изданных после 01.10.89.

Gazette of Patents and Trademarks

<http://www.hpo.hu/szkvlekerform.cgi>

Краткие библиографические данные о патентах Венгрии, изданных после 01.03.96. Информация на венгерском языке, приводится английский перевод названий патентов.

D'Agostini Organizzazione

<http://www.dagostini.it/patchclass/patchclass>

Неофициальное издание, содержащее списки классов и подклассов по Международной патентной классификации. Поисковое средство отсутствует. Все сведения даны на английском, немецком, французском, испанском и итальянском языках.

Степень информированности о патентах европейских стран может улучшиться во второй половине 1998 г., поскольку *European Patent Office* (<http://www.european-patent-office.org>) планирует открыть свободный доступ к своим ресурсам – первоначально к титульным страницам, а впоследствии и к полным текстам документов [5].

1. Patent Information from Online Databases. Karlsruhe, 1993.
2. Duberman J. // DATABASE. 1994. Vol.17, № 4. P.55.
3. Lambert N. // Ibid. 1996. Vol.19, №4. P.56.
4. Qi Y., Liu Y. // J. Chem. Inf. Comput. Sci. 1997. Vol.37, №6. P.1082.
5. Information World Review. 1998. №132. P.1.

Поступила в редакцию 03.03.98