

Особености при изграждане на автоматизирани информационни системи по геология на микрокомпютър

Е. Иванчев, М. Еревинова

Научноизследователски институт по полезни изкопаеми, 1505 София

E. Ivanchev, M. Erevinova — Specific features in the development of microcomputer information systems in geology. Input documents for automatized documental information systems (ADIS) for microcomputer in geology contain data on the bibliographic description, information on the content of the initial source and regional and geochronologic position of the studied object. They are written on the screen and do not contain a review. The information model of the initial source is based on the vocabulary of the polylingual tesauros on geology "Geoinform" supplemented mainly by regional terminology. The experimental tests show good results.

Геоложката информация представлява както основен продукт на научните изследвания и геологопроучвателните работи, така и един от главните ресурси на бъдещите дейности в тази област. Редица нейни особености (Иванчев и Кемилев, 1984; Стайков и Иванчев, 1985) правят въпросите по рационалното ѝ ползване особено остри и актуални. Най-важните и неотложни проблеми, които следва да се решат в това направление, са осигуряване на събирането, съхраняването и многократното използване на геоложката информация. На съвременния етап от развитието на електронноизчислителната техника най-подходящ метод за решаването на тези проблеми е изграждането на различни автоматизирани информационни системи (АИС).

Автоматизираните информационни системи се състоят от три основни компонента (Криницкий и др., 1982): физическа, която е носител на информацията; информационна, представляваща организирана по някакъв начин система от записи; функционална, включваща различни процедури по създаването, поддържането и търсенето в базите от данни.

В зависимост от задачите, които решават, АИС биват основно два типа — документални (АДИС) и фактографски (АФИС). Тази подялба отразява предимно формата на предоставяне на резултатите от информационното търсене — документ (обикновено вторичен) или даден списък от данни (факти), отразяващи конкретни характеристики на обекта на изследване. Според характера и степента на достъпност на първоизточниците, от които се набират данните, АДИС и АФИС биват по публикувани и по непубликувани документи.

Степента на удовлетворяване на постоянно растящите информационни потребности в областта на геологията у нас на настоящия етап е минимална. Това се обуславя главно от изоставане в създаването и целенасоченото натрупване на информационни масиви в паметта на ЕИМ. Бързият прогрес в производството на електронно-

изчислителна техника и тенденциите към все по-широко прилагане на персонални компютри са важна предпоставка за преодоляване на това изоставане. Налага се изработването на диференциран подход при проектирането и създаването на различни бази от данни. Като се има предвид, че ефективността от информационното обслужване се определя не само от релевантността, пълнотата и точността на намерената информация, но и от месторазположението на базата от данни, от нейната достъпност, времето и цената за обслужване на потребителите, то става очевидно, че за определени класове информационни задачи, използването на персонални компютри се превръща в оптимално решение. Такива задачи са създаването на тематично разпределени АДИС и специализирани АФИС, с помощта на които да се обслужват научни и производствени организации и звена, отличаващи се със специфика на своите обекти на изследване и тематика.

Преминаването към серийно производство на различни персонални микрокомпютри в нашата страна, ускореното разработване и внедряване на разнообразни програмни системи, в т. ч. и такива, които са подходящи за използване в АИС, решават проблемите, свързани с физическия и функционалния компонент на системата. Така към персоналните микрокомпютри „Правец-82“ е създадена програмна система ПФС (персонална файлова система), чиито главни функции — проектиране, добавяне, копиране, търсене, печат и отстраняване, и техните различни подфункции удовлетворяват основните изисквания към функционалния компонент.

Главният проблем при създаването на базите от данни по геология се състои в правилното изграждане на структурата и лингвистичното осигуряване на информационния им компонент. Този проблем е единствено от компетентността на специалистите в областта на геоложката информатика.

В последно време бяха извършени някои изследвания, свързани със създаването на АИС по геология на микрокомпютър. Разработени бяха проекти за АДИС по публикувана (Иванчев и др., 1985) и непубликувана (колектив под ръководството на М. Стайков) геоложка литература, както и АФИС по хидрогеохимия¹. Резултатите от опитното функциониране на АИС на базата на създадените масиви дават възможност за анализ и достигане до някои съществени изводи.

Структурата и съдържанието на входните документи на АДИС на микрокомпютър трябва да отговарят на някои основни изисквания. Преди всичко от гледна точка на икономия на машинна памет, бързо въвеждане на информацията и удобно ползуване на системата, всички данни за първичния документ се разполагат на един екран. Освен това входният документ на АДИС както по публикувана, така и по непубликувана информация съдържа три групи данни, разположени в съответни полета: библиографско описание, сведения за съдържанието на първоизточника, регионална и геохронологика привързаност на обекта на изследване (фиг. 1).

Част от библиографските данни — автор(и), заглавие, година се отнасят и до двата вида първични документи. Освен тях при публикуваните следва описание на първоизточника, извършено по съответни стандарти, а за непубликуваните в тази група се включват архивният номер, № на микрофиш, предприятието (организацията), извършило проучването. Архивният номер на непубликувания първоизточник (доклад, отчет и т. н.) е основен признак, осигуряващ съвместимостта на документалната АИС и различни бъдещи АФИС, които ползват данни от същия първоизточник.

Видът и начинът на подготовка на сведенията за съдържанието на първоизточника играят основна роля за достигане на необходимите пълнота и точност на резултатите от информационното търсене. При микрокомпютрите е неподходяща широко използваната практика за въвеждане на реферат поради непригодността на програмната система ПФС за ефективно търсене в свободен текст, необходимостта от голяма машинна памет и бавно въвеждане на информацията.

¹ Описанието на АФИС по хидрогеохимия е предмет на отделна публикация.

съответното направление. Например в дискетата, съдържаща инженерногеоложки документи, в това поле се въвеждат главните подразделения на инженерната геология — грунтознание, регионална инженерна геология и т. н. Ако описваният първоизточник е строго специализиран, може да се въведе още едно, йерархично по-ниско, ниво на поделяне на науката, например „теоретично грунтознание“.

При АДИС по непубликувана литература, изграждаща се с оглед на автоматизация на търсенето на документи в Геофонда на Комитета по геология, разпределянето на информацията върху различни дискети се извършва в съответствие с разделите, в които са описани първоизточниците. Така например докладите от геоложката картировка (раздел IV) се разполагат на отделна дискета. Специализираният характер на информацията в Геофонда обаче налага индексиранието на съдържанието на докладите да се концентрира главно върху обекта на изследване, използваните методи, етап на проучване, типа на запасите, като за всеки от тези показатели е разработен специален класификатор. Класификаторът е друга форма на представяне на езиковото осигуряване на системата. Неговото предназначение е да отрази формализирано всички възможни термини, които могат да се използват за попълване на дадено поле от машинната форма. Служи за еднотипно въвеждане на документите и запитванията с оглед на висока пълнота и точност при търсенето.

Обектът на изследване най-често е свързан с полезното изкопаемо и неговите различни форми на проявление. Така например металните полезни изкопаеми, които се разглеждат в доклада, се описват чрез химическия знак на съответния елемент. Останалите неметални, течни и газообразни суровини се описват, като се ползува специално разработен класификатор.

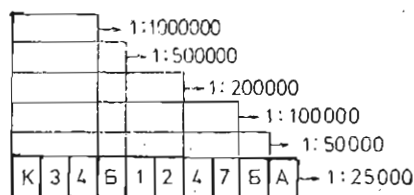
Регионалната и геохронологичната привързаност на обекта на изследване са друг основен аспект на значителна част от геоложките документи. Тази група от данни е особено важна при обработката на националната публикувана и непубликувана литература. Необходимо е различните АДИС да съдържат еднотипно подготвени данни за географската привързаност, местното название, окръга и стратиграфския код на обекта на изследване, което осигурява съвместимостта на системите.

Опитът показва, че най-добри резултати при търсенето се постигат, когато регионалното привързване се извършва в две полета на машинния документ (фиг. 1). В първото поле, ГЕОГ, в съответствие с разработен класификатор се указва към коя крупна географска единица се отнася районът, разглеждан от описвания документ. Така например Източни Родопи се нанася като РОДОПИ=И. Тази форма на въвеждане на информацията дава възможност при необходимост да се търсят данни и за цялата Родопска област. В такъв случай запитването се индексира с използване на маски като . . РОДОПИ. . . Полето НАЗВ се разглежда като допълнение и уточнение на географската привързаност. В него се нанасят най-важните местни географски названия — градове, села (съгласно Единен класификатор на населените места в НР България), реки, планини, езера, върхове и др. (съгласно специално разработен класификатор). След названието се указва за какъв обект става дума. Този начин на представяне изключва възможността за информационен шум, който би могъл да се получи, тъй като различни географски обекти имат еднакви названия: например Доспат=р (река), Доспат=яз (язовир), Доспат=г (град) или Доспатска=пл (планина), Доспатска=котл (котловина). Словосъчетанията, подобно на различни съществуващи АИС, също се свързват със знак за равенство. Например връх Кройдово тепе се записва КРОЙДОВО=ТЕПЕ=ВР.

Окръзите се нанасят в самостоятелно поле, като за тяхното отразяване се използва широко разпространения двубуквен мнемоничен код.

Стратиграфската привързаност на обекта на изследване се отразява в полето СТР, като се нанасят индексите (кодовете) на стратиграфските единици от ера до етаж съгласно хроностратиграфската таблица. Например горен триас се кодира като MZT3, юра — като MZJ, а сенон — като MZK2SN. Този начин на кодиране има голямо предимство, тъй като предоставя възможността за йерархично търсене на ин-

формацията. Нека например, ако запитването касае горен триас, то ще се индексира като „МЗТЗ...“. Ако се търсят данни за триаса въобще, то запитването ще бъде „МЗТ...“, при което ще излязат и документите, в които се съдържат данни освен за горния, но и за средния, и долния триас. В случай, че потребителят се интересува от мезо-



Фиг. 2. Код на картен лист
К-34/-Б-ХII/-47-Б-а

зоа като цяло, то запитването „МЗ..“ ще изведе информация за триаса, юрата и кредата с всичките им подразделения.

За АДИС по национална публикувана геоложка литература такава регионална и стратиграфска привързаност е напълно достатъчна. За непубликуваните разработки обаче се налага да се извърши и картографско привързване.

Това става в полето К.Л (картен лист) като се използва „Сборен лист за карти в М 1 : 1 000 000, 1 : 500 000...“. Попълва се стандартното обозначение на картния лист, който обхваща изцяло разглеждания район. Отново, подобно на стратиграфската привързаност, се използва възможността за йерархично търсене на информацията, като се въвеждат последователно обозначенията на картни листове в мащаб от 1 : 1 000 000 до 1 : 25 000 (фиг. 2). Ако районът попада в повече от три съседни картни листа от даден мащаб, тогава се попълва индексът на картния лист в по-малък мащаб, който ги включва. Полето се проектира така, че позволява въвеждане на обозначенията най-много на три картни листа в М 1 : 25 000. Римските цифри в обозначенията на картни листове в М 1 : 200 000 се заменят с арабски с оглед на икономия на място.

Следва да се подчертае, че за АДИС по публикувана литература показаните на фиг. 1 полета е необходимо да се проектират за всички геологически науки. Тяхната дължина обаче може да бъде различна в зависимост от конкретната научна област. В инженерногеоложките публикации например рядко се разглеждат много и различни хроностратиграфски единици, поради което полето СТР има малка дължина.

Така проектираните АДИС дават възможност за търсене на информация по най-разнообразни запитвания (табл. 1). Обикновено запитванията, свързани с библиографското описание, се свеждат до име на автора или източника. По-голяма част от въпросите се отнасят до съдържанието на документа и регионалната привързаност. Някои се отнасят само до съдържанието на документа, а други са свързани само с конкретен район от страната.

От показаните примерни въпроси се вижда, че обикновено търсенето става по текстово съответствие в дадено поле. Само при въпрос № 1 се извършва търсене и по числово условие. Между отделните търсещи полета се изпълнява функцията логическо умножение, т. е. необходимо е едновременно изпълнение на всички зададени условия (фиг. 3).

Резултатите от експерименталното изпробване на АДИС показват добрите възможности на микрокомпютрите и програмното им осигуряване за съхраняване и намиране на необходимата информация. За да се осигури съвместимост на тематично обособените подсистеми е необходимо да се използва общо лингвистично осигуря-

Таблица 1

Примерни запитвания към АДИС по публикувана национална геоложка литература на микрокомпютър

| № | Въпрос | Търсещо предписание | |
|---|---|-------------------------|--|
| | | поле | съдържание на полето |
| 1 | Публикации на проф. Демирев в периода 1980—1982 г. | АВТ Г | . . А. Демирев. . =80. . 82 |
| 2 | Публикации на проф. Демирев в годишника на ВИАС | АВТ ИЗТ | . . А. Демирев. . . Год. ВИАС . |
| 3 | Инженерногеоложки изследвания на Добруджанския въглищен басейн | НАЗВ | . Добрудж=ВЪГЛ=БАС.. |
| 4 | Корелационни зависимости между свойствата на глините | ОБ ПР | . Глина. . . Корелация. . |
| 5 | Математически методи за прогноза на инженерногеоложките условия | ОБ ПР МЕТ | . ИГ=условия. . . Прогноз. . . Математ. . |
| 6 | Изследване на якостните свойства на плиоценските глин от Софийската котловина | ОБ СВ ГЕОГ | . Глина. . . ЯК. . . Софийска=котл. . |
| 7 | Деформации на откосите на рудниците на Марица-изток | СТР ОБ ПР НАЗВ | . NZTRN2. . . Откос. . . Деформация. . . Маришки=БАС=И. . |

N: 12 АВТ: А.ДЕМИРЕВ.Р.ГЮРЧОВ

Г: 81

ЗАП: ОПЪРНА ЯКОСТ НА НЯКОИ ПЛИОЦЕНСКИ ГЛИНИ В УСЛОВИЯ НА СЛОЖНИ НАПРЕЖАТИ СЪСТ
ОИНИЕ

ИЗТ: ГОД. ВИАС Св.4 117-125

НАУКА: ГРУНТОЗН

ОБ: ДИСП=ГЛИНА АП=ЛАБ= ПОЛ=

ПР:

МЕТ: ИЗСЛ КОНСТР СРЪЗВ=ПН НАТИСК=ТРИОС

СВ: МЕХ=ЯК

ГЕОГ: СОФИЙСКА=КОТЛ

НАЗВ: СОФИИ=Г

СТР: NZTRN2

ОКР: СГ

N: 11 АВТ: А.ДЕМИРЕВ.Л.ИЛИЕВА

Г: 81

ЗАП: ИЗПОЛЗУВАНЕ НА НЯКОИ КОРЕЛАЦИИ МЕЖДУ ПОКАЗАТЕЛИТЕ НА ФИЗИКО-МЕХАНИЧНИТЕ С
ВОЙСТВА НА ГЛИНИ ПРИ ПРОГНОЗИРАНЕ НА ИНЖЕНЕРНО-ГЕОЛОЖКИТЕ УСЛОВИЯ ЗА СТРОИТЕЛСТВ
О

ИЗТ: ГОД. ВИАС Св.4 127-136

НАУКА: ГРУНТОЗН

ОБ: ИГ=УСЛОВИЯ ДИСП=ГЛИНА АП=ЛАБ= ПОЛ=

ПР:

МЕТ: МАТ=СТАТИСТ=КОРЕЛ ПРОГНОЗ

СВ: МЕХ=ЯК ФИЗ=ПЪПН ФИЗ-ХИМ=ПЛАСТ СЪСТ

ГЕОГ: СОФИЙСКА=КОТЛ

НАЗВ: СОФИИ=Г

СТР: NZTRN2

ОКР: СГ

Фиг. 3. Разпечатка на два документа, релевантни на въпрос 6

ване. МТ по геология може да се разглежда като подходяща основа за създаване на национален геоложки тезаурус, който да се издаде в необходимия тираж. Освен това входните документи следва да бъдат проектирани еднотипно, като съдържат посочените на фиг. 1 полета. Опростеният човеко-машинен диалог и системното отсъствие на цена на машинното време правят въвеждането на подобни системи икономически ефективно. Като се има предвид, че годишната национална публикувана продукция по геология е между 300—500 документа, което означава около три човекомесеца работа на квалифициран информатор за обработката и въвеждането на документите, то става очевидно, че създаването на такава АДИС е преди всичко организационен въпрос, който е от компетентността на национална организация по геология като БГД.

Л и т е р а т у р а.

Единен класификатор на населените места в НР България. 1982 — КЕССИ. 289 с.

Иванчев, Е., Вл. Кемилев. 1984. Изграждане на автоматизирана система по непубликувана инженерногеоложка и хидрогеоложка информация — *Строителство*, 12, 38—41.

Иванчев, Е., Т. Кехайов, М. Еревинова. 1985. Возможности применения персонального компьютера Правец-82 для информационного обеспечения геологических задач. — Сб. *Горнорудн. Пржибрам в науке и технике*, секция „Мат. методы в геол.“, VIII/Д4.

Криницкий, Н. А., Г. А. Миронов, Г. Д. Фролов. 1982. *Автоматизированные информационные системы*. М., Наука. 381 с.

Стайков, М., Е. Иванчев. 1987. Някои принципи за изграждане на автоматизирани информационни системи по геология у нас. — В: *Проучване и развитие на минерално-суровинната база на НРБ*. С., Техника (под печат).

(Постъпила на 14. IV. 1986 г.)