

22502^{+CD}
m. 2

Wiesław Gliński • Henryk Rybiński

TWORZENIE I UDOSTĘPNIANIE BAZ DANYCH W SIECI WWW

WYDAWNICTWO

SBP



NAUKA - DYDAKTYKA - PRAKTYKA



WIESŁAW GLIŃSKI jest adiunktem w Zakładzie Systemów Informacyjnych Instytutu Informacji Naukowej i Studiów Bibliologicznych na Uniwersytecie Warszawskim, gdzie pracuje od 1992 r.

W 1990 r. ukończył studia filozoficzne w zakresie logiki i metodologii nauk na ATK (Akademia Teologii Katolickiej, obecnie Uniwersytet Kard. Stefana Wyszyńskiego). Doktorat z zakresu informacji naukowej obronił na Uniwersytecie Warszawskim.

Prowadzi zajęcia na Uniwersytecie w zakresie technologii i systemów informacyjnych. Prowadził również gościnne wykłady m.in. na: Politechnice Warszawskiej, w Fachhochschule Hannover – Fachbereich Informations und Kommunikationswesen w Hannoverze (Niemcy), w SAXION Hogeschool IJsseland, Institut Informatica en Informatie-management w Deventer (Holandii), w Ministerstwie Gospodarki w Asmarze (Erytrea), w Library Association w Londynie (Wielka Brytania), w Kent State University w Kent, Ohio w Stanach Zjednoczonych.

Dr Wiesław Glišński był również laureatem wielu stypendiów, m.in. Fulbrighta, IMG Tempus Phare.

W swojej pracy naukowo-badawczej zajmuje się ontologiami, udostępnianiem baz danych w sieci WWW, zdalnym kształceniem, sieciami informacyjnymi, inteligentnymi systemami informacyjnymi. Był opiekunem kilkunastu prac licencjackich. Jest autorem bądź współautorem kilkudziesięciu prac oraz jednej książki z zakresu systemów informacyjnych.

Od końca lat dziewięćdziesiątych jako konsultant uczestniczy w wielu projektach dotyczących systemów informacyjnych dla UNIDO, FAO, PTC (Polskiej Telefonii Cyfrowej), Instytutu Inżynierii Informatycznej.



HENRYK RYBIŃSKI jest absolwentem Politechniki Warszawskiej. W latach 1975-1989 pracował w Instytucie Informacji Naukowej, Technicznej i Ekonomicznej. W latach 1980-1985 był ponadto wykładowcą w Instytucie Bibliotekoznawstwa i Informacji Naukowej. Od 1989 r. pracuje na Politechnice Warszawskiej. Od 1994 r. kieruje Zakładem Systemów Informacyjnych w Instytucie Informatyki Politechniki Warszawskiej. W 1993 r. otrzymał stanowisko profesora Politechniki Warszawskiej. W 2001 r. otrzymał tytuł naukowy profesora.

Jego działalność badawcza koncentruje się na teorii i praktyce systemów informacyjnych, w szczególności metodach reprezentacji wiedzy, zagadnieniach eksploracji danych, teorii zbiorów przybliżonych.

Jest zaangażowany w działalność dydaktyczną w zakresie zaawansowanych problemów baz danych, analizy i projektowania systemów informacyjnych, metod inteligentnych systemów informacyjnych. Był opiekunem ponad 40. prac magisterskich i 4. doktoratów. Jest autorem bądź współautorem ponad 80. prac w międzynarodowych czasopismach i na konferencjach oraz 4. książek.

Od połowy lat osiemdziesiątych jako ekspert ONZ uczestniczy w wielu projektach dotyczących systemów informacyjnych dla UNIDO, FAO, UNESCO, IFAD, WFP i innych.

TWORZENIE I UDOSTĘPNIANIE BAZ DANYCH W SIECI WWW

Polish Librarians Association
SCIENCE-DIDACTICS-PRACTICE

Wiesław Gliński • Henryk Rybiński

**DATABASE DESIGNING AND
PUBLISHING ON THE WEB**

The handbook fo librarians and information specialists



Warsaw 2004

Stowarzyszenie Bibliotekarzy Polskich
NAUKA-DYDAKTYKA-PRAKTYKA

Wiesław Gliński • Henryk Rybiński

TWORZENIE I UDOSTĘPNIANIE BAZ DANYCH W SIECI WWW

Podręcznik dla bibliotekarzy i dokumentalistów

WYDAWNICTWO

SBP



Warszawa 2004

Komitet Redakcyjny serii wydawniczej
<<NAUKA — DYDAKTYKA — PRAKTYKA>>

Marcin DRZEWIECKI (przewodniczący), Stanisław CZAJKA, Artur JAZDON,
Barbara SOSIŃSKA-KALATA, Danuta KONIECZNA, Krzysztof MIGOŃ, Mieczysław
MURASZKIEWICZ, Janusz NOWICKI (sekretarz), Joanna PAPUZIŃSKA-BEKSIĄK,
Wanda PINDŁOWA, Jadwiga SADOWSKA, Jan SÓJKA, Barbara STEFANIAK,
Elżbieta STEFAŃCZYK, Hanna TADEUSIEWICZ, Zbigniew ŻMIGRODZKI

**Publikacja dofinansowana przez Instytut Informacji Naukowej
i Studiów Bibliologicznych Uniwersytetu Warszawskiego**

Recenzent
Mieczysław MURASZKIEWICZ

Redakcja techniczna i korekta
Anna LIS

© Copyright by Stowarzyszenie Bibliotekarzy Polskich

ISBN 83-89316-20-X



CIP — Biblioteka Narodowa
Gliński, Wiesław

Tworzenie i udostępnianie baz danych w sieci WWW : podręcznik dla bibliotekarzy
i dokumentalistów / Wiesław Gliński. Henryk Rybiński ; Stowarzyszenie Bibliotekarzy
Polskich. - Warszawa : Wydaw. SBP, 2004. - (Nauka, Dydaktyka, Praktyka ; 68)

Wydawnictwo SBP. 00-335 Warszawa, ul. Konopczyńskiego 5/7, tel./fax 827-52-96

Wydanie I. Warszawa, 2004 r. Ark. wyd. 12,2. Ark. druk. 14,5
Łamanie: K. BRAWIŃSKI. Druk i oprawa: Z. P. Poligrafia
00-389 Warszawa, ul. Smulikowskiego 6/8, tel. 826-06-03

09.04; dep SBP 140.00,-

Spis treści

PRZEDMOWA	9
SKRÓTY I OZNACZENIA	11
1. WYBRANE ASPEKTY MODELOWANIA E- R	13
1.1. Wstęp	13
1.2. Ogólne podejście	14
1.3. Podstawowe pojęcia	15
1.4. Zależności semantyczne	16
1.5. Przykłady modelowania E-R	17
1.6. Ćwiczenia	25
2. SYSTEM CDS/ISIS – WYBRANE ASPEKTY	27
2.1. Wstęp	27
2.2. Struktura baz danych i architektura systemu	32
2.3. Modelowanie E-R w systemie CDS/ISIS	34
2.4. Udostępnianie baz ISIS w sieci WWW	36
2.5. Ćwiczenia	38
3. WWW-ISIS – OGÓLNY PRZEGLĄD SYSTEMU	41
3.1. Ogólny schemat funkcjonowania systemu	41
3.2. Serwery i przeglądarki WWW	43
3.3. Instalacja wersji DEMO systemu	45
3.4. Katalogi i zbiory instalacyjne	52
3.5. Podstawowe funkcje systemu	55
3.6. Interfejs wprowadzania danych	59
3.7. Ćwiczenia	67
4. WYSZUKIWANIE W WWW-ISIS	69
4.1. Wstęp	69
4.2. Plan działania	70
4.3. Podstawowe kroki tworzenia systemu wyszukiwawczego WWW-ISIS	73
4.4. Metoda skrócona implementacji systemu WWW-ISIS	75
4.5. Standardowy sposób implementacji systemu	84
4.6. Modyfikacja zbiorów HTML oraz zbiorów HTP	107
4.7. Udoskonalanie interfejsu	109
4.8. Ćwiczenia	120
5. WPROWADZANIE DANYCH W WWW-ISIS	123
5.1. Plan działania	123
5.2. Wstępna faza projektowania bazy	127
5.3. Implementacja systemu	131
5.4. Definiowanie arkusza wprowadzania danych	132
5.5. Narzędzia sprawdzania poprawności	149
5.6. Funkcje ON_NEW i ON_UPDATE	155

5.7. Prawa dostępu, grupy użytkowników	158
5.8. Inne aspekty interfejsu	159
5.9. Poczta elektroniczna, strony kodowe	170
5.10. Zbiory HTP i konfiguracyjne	173
5.11. Ćwiczenia	174
ANEKSY	
Aneks A: Konfiguracja – isis3w.exe.cfg	178
Aneks B: Elementy systemu CDS/ISIS	184
Aneks C: Konwersja danych do formatu ISO2709	205
Aneks D: Zależności konfiguracyjne	211
Aneks E: Środowisko sieci	212
Aneks F: Tablice FDT i FST	224
LITERATURA	228
SPIS RYSUNKÓW, TABEL I TABLIC	230
SKOROWIDZ WAŻNIEJSZYCH TERMINÓW	233

Table of contents

PREFACE	9
ABBREVIATIONS AND ICONS	11
1. SELECTED ISSUES OF E-R MODELING	13
1.1. Introduction	13
1.2. General overview	14
1.3. Basic notions	15
1.4. Semantic dependencies	16
1.5. Examples of E-R modeling	17
1.6. Exercises	25
2. CDS/ISIS SYSTEM – SELECTED ISSUES	27
2.1. Introduction	27
2.2. Database structure and system architecture	32
2.3. E-R modeling in CDS/ISIS	34
2.4. Publishing ISIS databases on the Web	36
2.5. Exercises	38
3. WWW-ISIS – GENERAL OVERVIEW	41
3.1. General model of the system	41
3.2. WWW servers and browsers	43
3.3. Installing DEMO version of the system	45
3.4. Folders and system files	52
3.5. Basic functions of the system	55
3.6. Data entry interface	59
3.7. Exercises	67
4. SEARCHING IN WWW-ISIS	69
4.1. Introduction	69
4.2. Preparing a work plan	70
4.3. Basic steps towards implementing search interface in WWW-ISIS	73
4.4. Fast track of implementing the system	75
4.5. Standard track of implementing the search interface	84
4.6. Modifications of HTML and HTP files	107
4.7. System improvement	109
4.8. Exercises	120
5. DATA ENTRY IN WWW-ISIS	123
5.1. Work plan	123
5.2. Preliminary steps towards designing the database	127
5.3. System implementation	131
5.4. Defining data entry worksheets	132
5.5. Validation tools	149
5.6. ON_NEW and ON_UPDATE function	155

5.7. Access authorization	158
5.8. Other aspects of the interface	159
5.9. Mailing tools, national coding pages	170
5.10. Modifications of HTML and HTP files	173
5.11. Exercises	174
APPENDICES	
APPENDIX A: isis3w.exe.cfg configuration	178
APPENDIX B: Basics on CDS/ISIS system	184
APPENDIX C: Converting data into ISO2709	205
APPENDIX D: Configuration dependencies	211
APPENDIX E: Network environment	212
APPENDIX F: FDT and FST	224
BIBLIOGRAPHY	228
INDEX OF TABLES AND FIGURES	230
INDEX	233

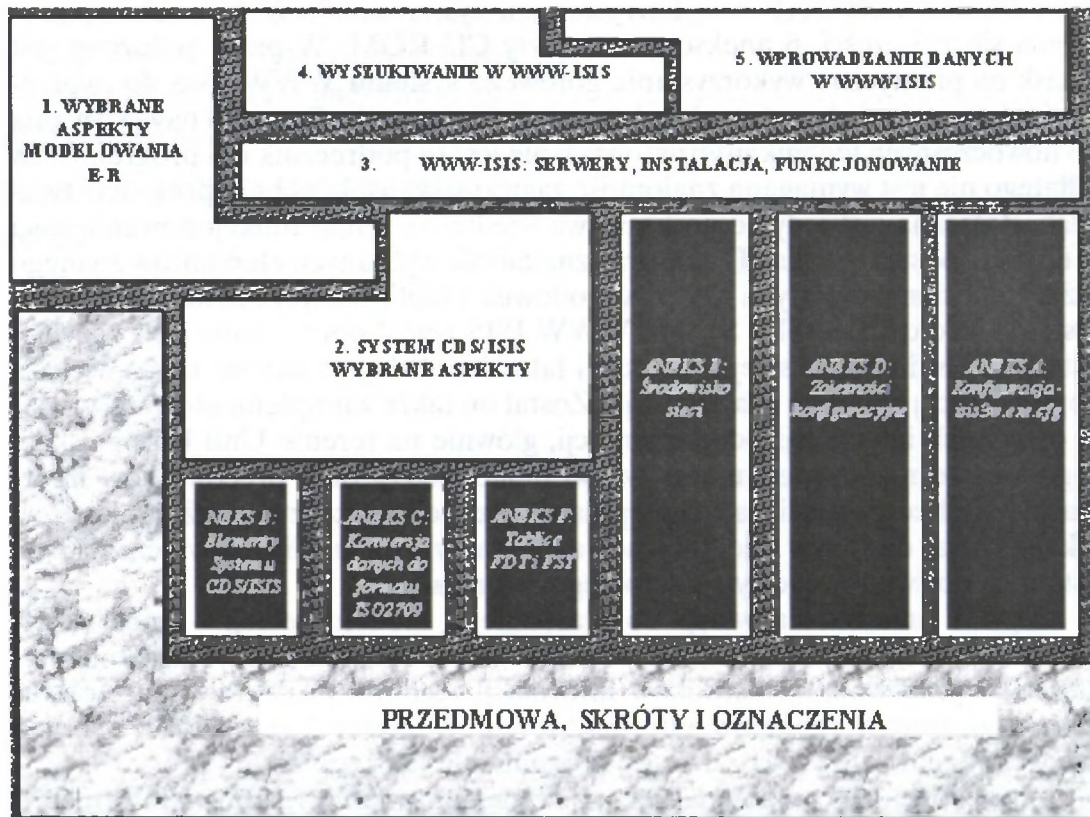
Przedmowa

Przedstawiamy Państwu nowy podręcznik poświęcony systemowi WWW-ISIS. Jest on dedykowany studentom wydziałów bibliotekoznawstwa, bibliotekarzom oraz pracownikom informacji naukowo-technicznej. Praca dotyczy tworzenia, utrzymywania internetowych oraz intranetowych baz danych w systemie Windows 95/98/NT/2000/XP z wykorzystaniem systemu WWW-ISIS. Podręcznik składa się z 5 części, 6 aneksów oraz płyty CD-ROM. W pracy położony jest nacisk na praktyczne wykorzystanie gotowego systemu WWW-ISIS do tworzenia i udostępniania baz danych w Internecie i intranecie. Pomimo nawiązywania do nowoczesnych technik internetowych nie jest to podręcznik dla programistów i dlatego nie jest wymagana znajomość zaawansowanych technik programistycznych. Zalecana jest jedynie podstawowa wiedza na temat funkcjonowania sieci Internet, podstaw języka HTML oraz znajomość wybranych elementów znanego już od lat osiemdziesiątych XX w. w środowisku bibliotekarzy i dokumentalistów systemu Micro-CDS/ISIS. System WWW-ISIS został opracowany przez specjalistów z Polski i dzięki swej prostocie i łatwości obsługi cieszy się coraz większą popularnością w Polsce i za granicą*. Został on także zaimplementowany w wielu ośrodkach informacji i dokumentacji, głównie na terenie Unii Europejskiej. Jego szersze rozpowszechnienie wśród bibliotekarzy oraz pracowników informacji w Polsce pozwoli na udostępnianie niewielkim kosztem już istniejących lokalnych baz danych w sieci Internet, a tym samym da możliwość włączenia tych ostatnich do światowego systemu obiegu informacji.

Część 1 przedstawia wstęp do klasycznego modelu relacji – encji (E-R) Chenna, przy czym jest ona opracowana w ten sposób, aby przybliżyć bibliotekarzom i pracownikom ośrodków informacji techniki projektowania struktur danych złożonych systemów informacyjnych. W części 2 omówiono wybrane aspekty systemu ISIS konieczne do zrozumienia dalszych części pracy (3, 4, 5) oraz niezbędne do w pełni świadomego implementowania systemów informacyjnych za pomocą WWW-ISIS; omówione są pokrótce wybrane systemy udostępniania baz ISIS w Internecie/Intranecie. Część 3 analizuje kwestie instalacyjne pakietu WWW-ISIS wraz z przykładową aplikacją (bazą) DEMO. Część 4 opisuje system wyszukiwawczy WWW-ISIS, jego konfigurację oraz utrzymywanie. Część 5 przedstawia system wprowadzania danych WWW-ISIS, jego instalację, konfigurację oraz utrzymywanie. Całość kończą spis literatury oraz 6 aneksów. Płyta CD-ROM zawiera oprogramowanie wybranych serwerów WWW, pełną wersję systemu WWW-ISIS z przykładową aplikacją DEMO, dokumentację oraz zbiory samouczków w języku angielskim.

* Np.: <http://www.oefse.at/oefse> (serwer Österreichische Forschungsförderung für Entwicklungshilfe), <http://faolex.fao.org/faolex> (witryna FAO – Food And Agriculture Organization of the United Nations) – bazy prawnicze), <http://faolex.fao.org/fishery> (witryna FAO), <http://www4.fao.org/faobib> (witryna FAO – katalog biblioteczny), <http://library.iccrom.org/libris> biblioteka ICCROM – International Centre for the Study of the Preservation and Restoration of Cultural Property, <http://www.gtz.de/uvp/english/uvsonline-e.htm> (Deutsche Gesellschaft für Technische Zusammenarbeit GTZ GMBH).

Każda z części stanowi pewną niezależną całość. Niemniej w celu właściwego zrozumienia poszczególnych rozdziałów konieczne jest zapoznanie się z pozostałymi, co przedstawia ogólny plan książki (rys. 1). Każdy prostokąt (wielokąt) reprezentujący rozdział (podrozdział i/lub aneks) wspiera się na prostokątach, wielokątach leżących niżej, choć wielkość prostokątów/wielokątów nie odzwierciedla objętości danego rozdziału. Tak więc nie można zrozumieć rozdziału 4 lub 5 bez zapoznania się z rozdziałem 3. Jednak rozdział 1 można czytać niezależnie od rozdziałów 3, 4 lub 5. Aneksy nie są wymagane do zrozumienia poszczególnych części podręcznika, pod warunkiem, że czytelnik jest już zaznajomiony z systemem CDS/ISIS i potrafi skonfigurować sieć oraz serwer WWW.



Rys. 1. Ogólny plan podręcznika

Podręcznik ten jest zalecany dla studentów bibliotekoznawstwa i informacji naukowej, m.in. dla przedmiotów dotyczących baz danych, opisów bibliograficznych, dokumentacji elektronicznej, edytorstwa, informacji biznesowej, sieci komputerowych i Internetu, organizacji i zarządzania bibliotekami, systemów informacyjnych, zautomatyzowanych systemów bibliotecznych i informacyjnych.

Skróty i oznaczenia

Dla ułatwienia przyjęto w pracy szereg skrótów i oznaczeń ikonograficznych, których celem jest szybkie zorientowanie się czytelnika w treści czytanego tekstu. Tak więc ćwiczeniom towarzyszą następujące piktogramy przedstawione w tabl. 1.

Tablica 1

Ważniejsze skróty i oznaczenia przyjęte w książce

Ćwiczenie	
Ćwiczenia nie wymagające dostępu do komputera	
Ćwiczenia wymagające pracy z komputerem	
Ćwiczenia wymagające komputera zaopatrzonego w nagrywarkę CD-R	
Dotyczy wyłącznie systemu wyszukiwania danych WWW-ISIS (omówione głównie w rozdz. 4)	
Dotyczy wyłącznie systemu wprowadzania danych WWW-ISIS (omówione głównie w rozdz. 5)	
Działania (ćwiczenia) wymagające pracy z siecią i/lub serwerów WWW	
Przykład	
„Uwaga”, ważne zagadnienie, które należy zapamiętać	
Przypomnienie ważnej kwestii	
Dotyczy adresu URL widocznego w oknie przeglądarki (połączenie hipertekstowe)	
Wskazówka	
Założenie	
Ostrzeżenie przed problemami i wyjaśnienie, jak ich uniknąć	
Pytanie, na które należy udzielić odpowiedzi tak lub nie	
Oznacza szybką, choć nie zawsze najłatwiejszą metodę implementacji systemu	
Kolejny krok związany z wykonywaniem jakiegoś złożonego procesu	
Aby ułatwić czytelnikowi zorientowanie się, jakiego rodzaju umiejętności wymagane są w przypadku wykonania jakiegoś procesu lub ćwiczenia wprowadzono dodatkowo odpowiednie oznaczenia literowe:	
Do wykonania ćwiczenia konieczna jest znajomość modelowania przy wykorzystaniu relacji i encji	E-R

Do wykonania ćwiczenia lub kolejnego kroku w złożonym procesie wymagana jest znajomość języka HTML	<HTML>
Do wykonania ćwiczenia lub kolejnego kroku w złożonym procesie wymagana jest znajomość technik tworzenia zbiorów tablicy wyboru pól (tzw. FST – File Selection Table) w CDS/ISIS	CDS/ISIS FST
Do wykonania ćwiczenia lub kolejnego kroku w złożonym procesie wymagana jest znajomość technik eksportu/importu danych w CDS/ISIS	CDS/ISIS EXPORT/IMPORT
Do wykonania ćwiczenia lub kolejnego kroku w złożonym procesie wymagana jest znajomość technik charakterystycznych dla pakietu WWW-ISIS	WWW-ISIS
Do wykonania ćwiczenia lub kolejnego kroku w złożonym procesie wymagana jest znajomość języka formatowania CDS/ISIS	CDS/ISIS formaty

1. Wybrane aspekty modelowania E-R

1.1. Wstęp

Peter Chenn przedstawił swój model w późnych latach siedemdziesiątych. Podobnie jak inni teoretycy (Hammer oraz McLeod [HAM78]) odwoływał się do modelu semantycznego. Nowością w przypadku Chenna było odmienne podejście do zagadnienia danych. Model danych miał odwzorowywać rzeczywistość w najbardziej intuicyjny sposób, ale bez zwracania uwagi, w jaki sposób dany model ma być zaimplementowany w pamięci komputera. Swego czasu było to podejście całkiem rewolucyjne, pozwalające uwolnić się projektantom systemów baz danych od ograniczeń istniejących wówczas modeli reprezentacji danych [DAT00], dawało też możliwość modelowania rzeczywistości niezależnie od ostatecznego wyboru co do reprezentacji fizycznej danych [MURO93]. Pozwalało też wykrywać ewentualne nieprawidłowości modelu przed jego zaimplementowaniem. Podstawowa idea polegała na tym, że wszystko może być przedstawione w postaci pierwotnych abstrakcyjnych obiektów, które określane zostały mianem encji. Encje te mają pewne cechy lub jak wolimy atrybuty. Ponadto encje są powiązane między sobą przez związki. Związki te nazywamy relacjami. Motywacja badań nad modelami semantycznymi była następująca. Na ogół systemy baz danych – relacyjne lub inne – dysponują bardzo ograniczonym „rozumieniem” znaczenia danych reprezentowanych w bazie. Przeważnie rozumieją one pewne proste dane atomowe, bądź niektóre proste węzły integralności (reguły sprawdzające poprawność, czy dokładność danych w bazie) odnoszące się do tych danych. Jednak jest to bardzo ograniczone rozumienie danych (wszelka bardziej złożona interpretacja zależy od człowieka). Marzeniem projektantów było i jest, aby systemy mogły „rozumieć” nieco więcej. Dzięki temu mogłyby bardziej inteligentnie odpowiadać na działania użytkownika, a może nawet obsługiwać zaawansowane (tj. wyższego poziomu) interfejsy. Chcielibyśmy, aby np. jakiś język wyszukiwawczy rozumiał, że ciężary części (CZEŚCI), wielkości dostaw (DOSTAWY), chociaż wyrażone za pomocą liczb są wielkościami różnego rodzaju, tj. że są wielkościami znaczeniowo różnymi. Dzięki temu „rozumieniu”, żądanie połączenia (złączenia) dwóch tabel: CZEŚCI i DOSTAW na podstawie zgodnych ciężarów i ilości mogłoby być zakwestionowane albo nawet całkiem odrzucone. Według twórcy relacyjnych baz danych Codda rozpoznawanie danych jest ciągle niekończącym się procesem. Modelowanie semantyczne znane jest także pod innymi nazwami: modelowanie danych (data modeling), modelowanie związków encji (entity/relationship modeling), modelowanie encji (entity modeling). Niektórzy [DAT00] preferują termin modelowanie semantyczne z następujących powodów: termin modelowanie danych może kolidować z uznanym terminem „model danych” oznaczającym formalny system, złożony z obiektów, reguł integralności i operatorów, ponadto utrwała on błędny pogląd jakoby modelowanie danych dotyczyło jedynie struktury danych. Modelowanie związków encji zawęża rozumienie modelowania semantycznego sugerując tylko jeden rodzaj podejścia do problemu, podobnie jak termin modelowanie związków encji.

Zdaniem Date’a zasady modelowania semantycznego mogą stanowić użyteczną pomoc w systematycznym projektowaniu baz danych, nawet, jeśli dany System Zarządzania Bazą Danych nie wspiera takiego podejścia. Metody projektowania

oparte na zasadach modelowania semantycznego nazywa się często modelami „z góry na dół” [DAT00] ponieważ zaczynają proces projektowania od wysokiego poziomu abstrakcji obiektów świata rzeczywistego (tzw. encje itp.), a kończą na stosunkowo niskim poziomie abstrakcji w określonym projekcie bazy danych.

W celu ustalenia relacji między encjami w przykładowym modelu, należy określić jakie własności opisywanego świata nas interesują z punktu widzenia realizowanego systemu i jakie są ogólne związki pomiędzy opisywanymi obiektami. Jeżeli na przykład realizujemy system informacyjny wspomagający działanie pewnej instytucji, musimy ustalić, jakie obiekty (encje) będzie opisywać system, uzmysłwić sobie zasady, wedle których funkcjonuje dana instytucja (firma). Na przykład, jeżeli budujemy system o pracownikach i ich zaangażowaniu w realizowane projekty, aby mieć wymierne korzyści i móc w ogóle korzystać z przyszłej bazy danych, należy wybrać takie elementy opisu pracowników i projektów, które są istotne dla działania systemu. Musimy też zadbać o to, aby identyfikatory dowolnego pracownika, i/lub projektu oraz wydziału, w którym jest zatrudniony, były niepowtarzalne.

„Jak połączyć ze sobą pracowników, projekty oraz wydziały tak, aby zminimalizować powtarzalność danych a jednocześnie zwiększyć możliwości wyszukiwawcze?”.

Odpowiedzi na te i inne podobne pytania tworzą reguły, wedle których funkcjonuje (lub powinna funkcjonować) dana instytucja.

Definicje atrybutów wraz z określeniem relacji między encjami mogą posłużyć do określenia reguł, według których będzie funkcjonowała (lub funkcjonuje) dana instytucja. Na przykład można zdefiniować dopuszczalny przedział płac dla wszystkich pracowników lub tylko wybranej grupy. W ten sposób określamy domenę wszystkich dostępnych wartości dla danego atrybutu (w naszym przypadku atrybutu „Pensja”).

1.2. Ogólne podejście

Za Date'em ogólne podejście do procesu projektowania E-R można przedstawić za pomocą czterech kroków ([DAT00]).

W pierwszym kroku próbujemy w sposób nieformalny określić zbiór pojęć, użytecznych w dalszym procesie modelowania. Bazując na tym **intuicyjnym** pojęciu encji zakładamy, że świat składa się z encji, chociaż z góry nie ustalamy precyzyjnie, co jest, a co nie jest encją. Staramy się pogrupować encje w **typy encji**. Np. wszyscy wykładowcy na uczelni są instancjami ogólnego typu encji PRACOWNICY, wszystkie miesięczniki instancjami ogólnego typu encji CZASOPISMA. Oczywiście zaletą takiego grupowania jest to, że wszystkie encje danego typu będą miały pewne wspólne atrybuty (np. ubezpieczenie społeczne, pensje, ISSN, miejsce wydania itd.) oraz własności, co daje oczywiste oszczędności w reprezentacji. W szczególności można też przyjąć, że każda encja ma szczególną własność służącą do jej identyfikacji, tj. tzw. **tożsamość**. Encje mogą być połączone z innymi encjami za pomocą **związków**. Również związki pomiędzy encjami mogą charakteryzować się pewnymi własnościami. Rzecz jasna (jak zauważył Date), wspomniane terminy nie mają ścisłej formalnej definicji.

W kolejnym kroku próbujemy ustalić zbiór obiektów **symbolicznych** (formalnych), których można by użyć do reprezentowania powyższych pojęć.

W przedostatnim kroku określamy zbiór formalnych **reguł integralności** związanych z tymi formalnymi obiektami. Na końcu określamy zbiór **operatorów** do operowania nimi.

W dalszej części rozdziału ograniczymy się jedynie do nieformalnego przedstawienia podstawowych obiektów modelu E-R.

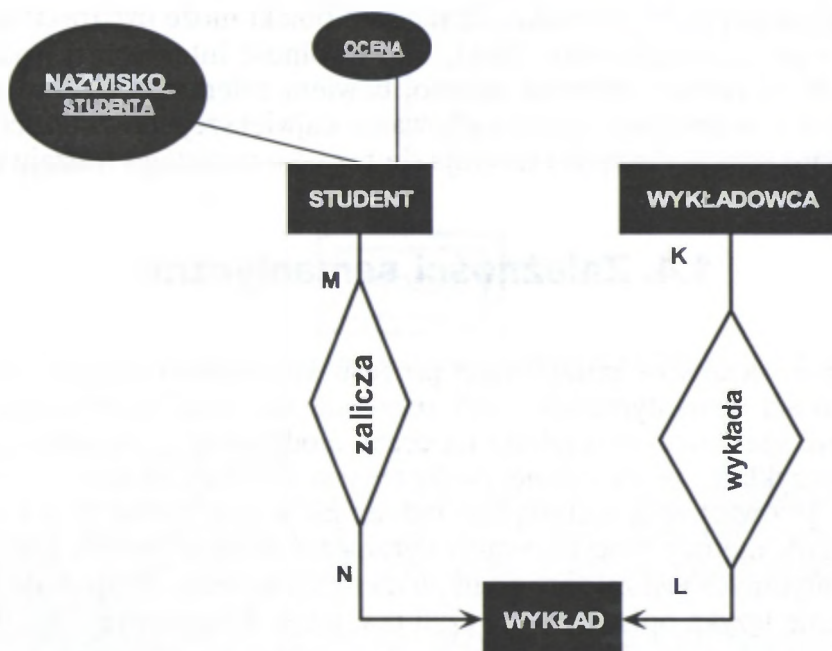
1.3. Podstawowe pojęcia

Jak już zostało wcześniej powiedziane, podstawowymi pojęciami modelu E-R jest obiekt i relacja. Struktura danych zorganizowana zgodnie z modelem E-R jest przedstawiana w formie graficznej jako diagram (rys. 2): prostokąty reprezentują obiekty, a ich typy reprezentowane są przez etykiety, zaś relacje reprezentowane są przez romby.



Rys. 2. Diagram E-R: „Pracownik pracuje nad projektem”

W rombikach umieszczona jest nazwa typu relacji, jak np.: **STUDENT**, **WYKŁAD**, **WYKŁADOWCA** oraz związki między nimi **ZALICZA**, **WYKŁADA**. Oba związki są wieloznaczne, rodzaje związków między typami obiektów są przedstawiane w diagramach w formie etykiet przy węzłach reprezentujących obiekty. W podanym niżej przykładzie wieloznaczność związku **STUDENT**, **WYKŁAD** przedstawiona jest przez etykiety **M** oraz **N**.



Rys. 3. Diagram E-R: „Student zalicza wykład”

W modelu E-R pojęcie relacji zarezerwowane jest jedynie do określania związków między obiektami. Zaś pojęcie atrybutu stosuje się do opisanie obiektów za pomocą pewnych prostych wartości. Pojęcie relacji ma tutaj zupełnie inne znaczenie niż w przypadku relacyjnego modelu Codd [MUR93], bowiem relacje w modelu E-R łączą złożone obiekty. Z formalnego punktu widzenia atrybut jest odwzorowaniem pomiędzy zbiorem obiektów (lub relacją), a zbiorem wartości.

Ponieważ ustalanie encji i relacji zależy w dużej mierze od intuicji oraz konkretnego wycinka modelowanej rzeczywistości przyjrzyjmy się nieformalnemu określeniu encji i relacji (wg [DAT00] tab. 1).

Tabela 1

Przykłady określenia podstawowych pojęć E-R

Pojęcie	Nieformalna definicja	Przykłady
ENCJA	Rozróżnialny obiekt	Dostawca, część, dostawa, student, pracownik, wydział, osoba, utwór, publikacja, wydawnictwo, artykuł, czasopismo, zamówienie itd.
WŁASNOŚĆ	Część informacji opisującej encje	Numer dostawcy, Numer indeksu, wielkość dostaw, Sygnatura książki, wydawnictwo, dział pracownika, wzrost osoby, kolor oczu, data zamówienia
ZWIĄZEK (RELACJA)	Encja służąca do połączenia dwóch lub więcej encji	Dostawca (dostawca-część), Zatrudnienie (pracownik-wydział), Publikacja (artykuł-czasopismo)
PODTYP	Encja typu Y jest podtypem encji typu X wtedy i tylko wtedy, gdy każdy Y jest też X	Pracownik jest podtypem osoby, Wykładowca jest podtypem pracownika, Rozdział jest podtypem książki

Z powyższego przykładu widać, że ten sam obiekt może być traktowany jako własność, a przez innych jako obiekt. Ta dowolność interpretacji (patrz przykłady w tab. 1) została dobrana celowo, bowiem celem modelowania semantycznego jest w miarę możliwości zachowanie największej *elastyczności interpretacji*. Ponadto relacje (związki) traktuje się też jako swoistego rodzaju encje.

1.4. Zależności semantyczne

Jednym z warunków pomyślnego procesu konceptualizacji jest określenie tzw. zależności semantycznych, czyli ustalenie własności tych danych, które powinny być spełnione ze względu na cechy modelowanego wycinka rzeczywistości. Na przykład, czy wiek danej osoby nie jest mniejszy od zera, czy sygnatura książki nie jest wartością ujemną itp. Jednak jak widać, model E-R nie określa automatycznie metody specyfikowania ograniczeń semantycznych. Język ograniczeń semantycznych jest zależny od implementacji systemu. Proponuje się w tym celu ustalenie języka opisu danych, czyli tzw. język deklaratywny, np. oparty na rachunku predykatów. Przykładowy język bazujący na rachunku predykatów znajdzie czytelnik w pozycji [MUR93].

1.4.1. Zależności semantyczne i diagramy E-R

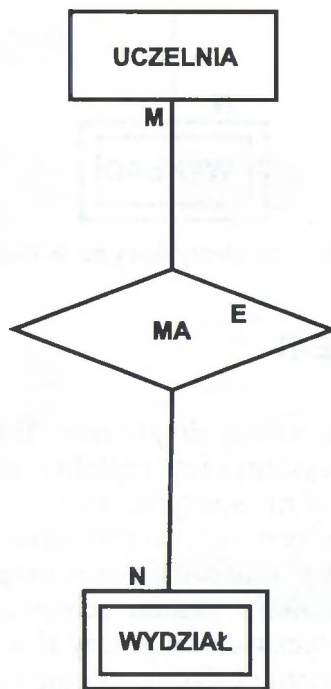
Mimo braku wyraźnego określenia zależności semantycznych, diagramy modelu E-R dają pewne możliwości określenia zależności semantycznych.

Zależność egzystencjalna

Przykładem może być tzw. zależność istnienia jednych obiektów od drugih, np. istnienie obiektu WYDZIAŁ jest zależne od obiektu UCZELNIA, co możemy zapisać w sposób nieco sformalizowany:

$$\text{WYDZIAŁY_UCZELNI} \subseteq \text{UCZELNIA} \times \text{WYDZIAŁ}$$

W powyższej relacji istnienie obiektu WYDZIAŁ jest zależne od istnienia obiektów typu UCZELNIA. Jeśli bowiem usuniemy z naszej bazy obiekty typu UCZELNIA, to również winny być usunięte wszystkie obiekty typu WYDZIAŁ. W diagramach E-R zaznacza się przeważnie ten fakt przez rysowanie podwójną linią obiektu zależnego i oznaczania relacji dodatkową literą np. **E** (*zależność egzystencjalna*).



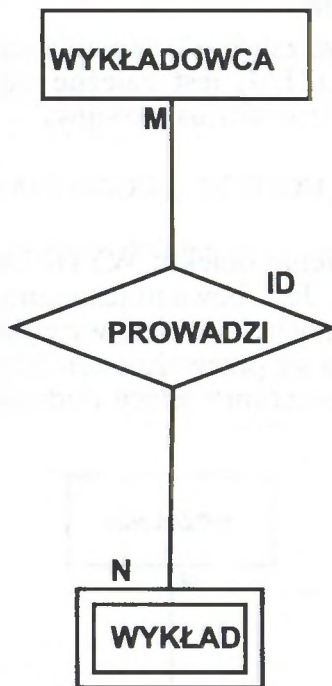
Rys. 4. Diagram E-R: Przykład zależności egzystencjalnej WYDZIAŁY_UCZELNI

Zależność identyfikacyjna

Innym przypadkiem ograniczeń semantycznych jest identyfikowalność obiektu przez jego powiązania, czyli gdy obiekt nie może być jednoznacznie identyfikowalny przez zbiór swoich atrybutów, a jedynie przez swoje związki z innymi obiektami.

Jeśli znowu odwołamy się do naszego przykładu z życia uniwersyteckiego, to w przykładowej bazie danych możemy zaobserwować istnienie obiektów WYKŁAD, które ze względu na swe atrybuty są nierozróżnialne, są natomiast

rozdzielalne przez swe związki z obiektami typu **WYKŁADOWCA**. Mówimy wówczas, że typ **WYKŁAD** jest w związku zależności identyfikacyjnej z typem **WYKŁADOWCA**. Zależność taką można reprezentować w podobny sposób, jak w rozważanym warunku istnienia z tą tylko różnicą, że kod zależności zapisujemy jako **ID** (zależność *Identyfikacyjna*, rys. 5.)



Rys. 5. Diagram E-R: Zależność identyfikacyjna WYKŁADOWCA_WYKŁAD

1.4.2. Zalety diagramów E-R

Podkreśla się, że istotną zaletą diagramów E-R jest ich czytelność, rola poszczególnych składowych systemu jest czytelna, gdyż wynika to bezpośrednio z relacji. Należy jednak rozróżnić specyfikowanie schematu danych (specyfikacje obiektów i powiązań między nimi), od zależności semantycznych.

To pierwsze jest w pełni wykonywalne przez diagramy i może znaleźć szereg praktycznych zastosowań w definiowaniu schematów danych. W przypadku określania zależności semantycznych diagramy E-R są co najwyżej narzędziem pomocniczym i dlatego konieczne było określenie odpowiedniego języka specyfikowania ograniczeń semantycznych.

1.5. Przykłady modelowania E-R

Poniżej zostanie omówionych kilka dowolnie wybranych przykładów zastosowania modelowania E-R. Ponieważ podręcznik ten jest przeznaczony dla bibliotekarzy i pracowników informacji, to szczególną uwagę zajmują w tym podrozdziale przykłady odwołujące się do tych wycinków modelowanego świata, które dotyczą bibliotek bądź baz bibliograficznych.

1.5.1. Przykład Relacji Student Zalicza Wykład

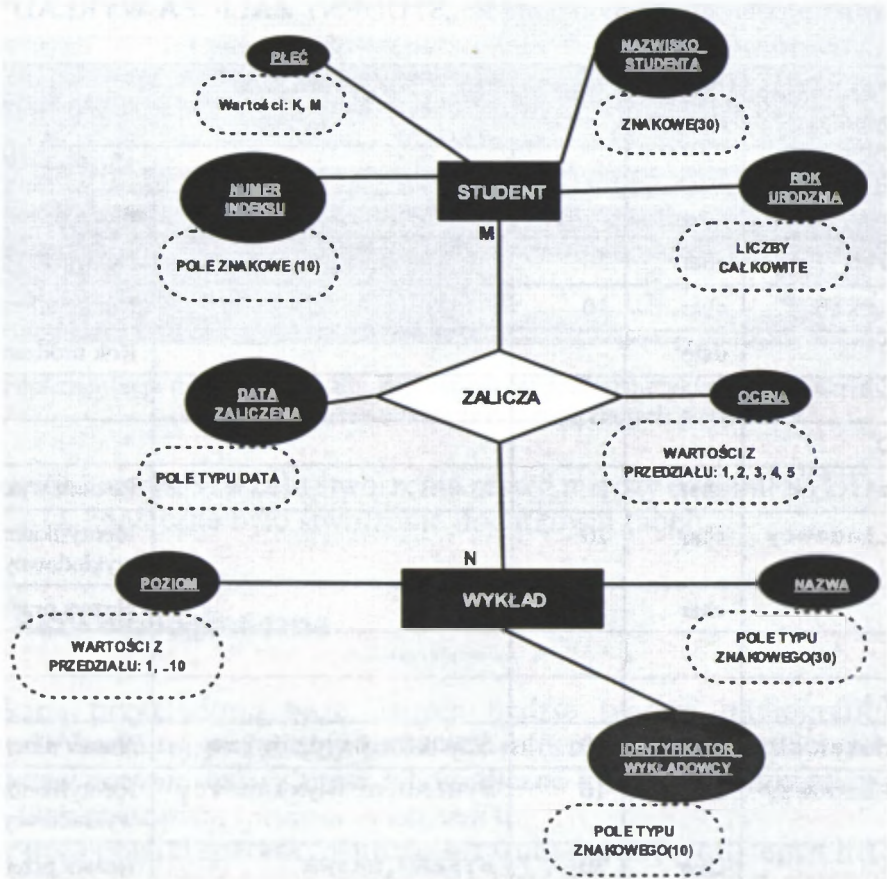
Przyjrzyjmy się простemu przykładowi modelowania E-R w przypadku studentów zaliczających wykłady na uczelni.



Rys. 6. Model E-R: „Student zalicza wykład” (skrócony)

Poszczególnym encjom przypisujemy atrybuty, np. NAZWISKO_STUDENTA, ROK_URODZENIA, itp. Atrybuty możemy też przypisać relacjom, np. relacji ZALICZA przypisujemy atrybut OCENA.

Zobaczmy, jakie atrybuty towarzyszą naszej przykładowej relacji.



Rys. 7. Diagram E-R: „Student zalicza wykład” (pełny)

Związywanie atrybutu OCENA z relacją ZALICZA pozwala uniknąć tworzenia sztucznego typu dla zapisu ocen studentów.

W tabeli 2 przedstawiona jest dokładna struktura danych w przykładowym modelu E-R dla relacji STUDENT ZALICZA WYKŁAD.

Poniżej, dla jasności, przedstawiliśmy ten sam model z opisaniem atrybutów. Niektóre z nich są kluczami podstawowymi tzn. jednoznacznie identyfikują encje, np. nr_indeksu identyfikuje encję STUDENT. W przypadku encji WYKŁAD wybrano jako klucz atrybut nr_wykładowcy.



Rys. 8. Diagram E-R: dla relacji „Student zalicza wykład”

W tabeli 2 przedstawiono strukturę danych z wyszczególnieniem klucza relacji.

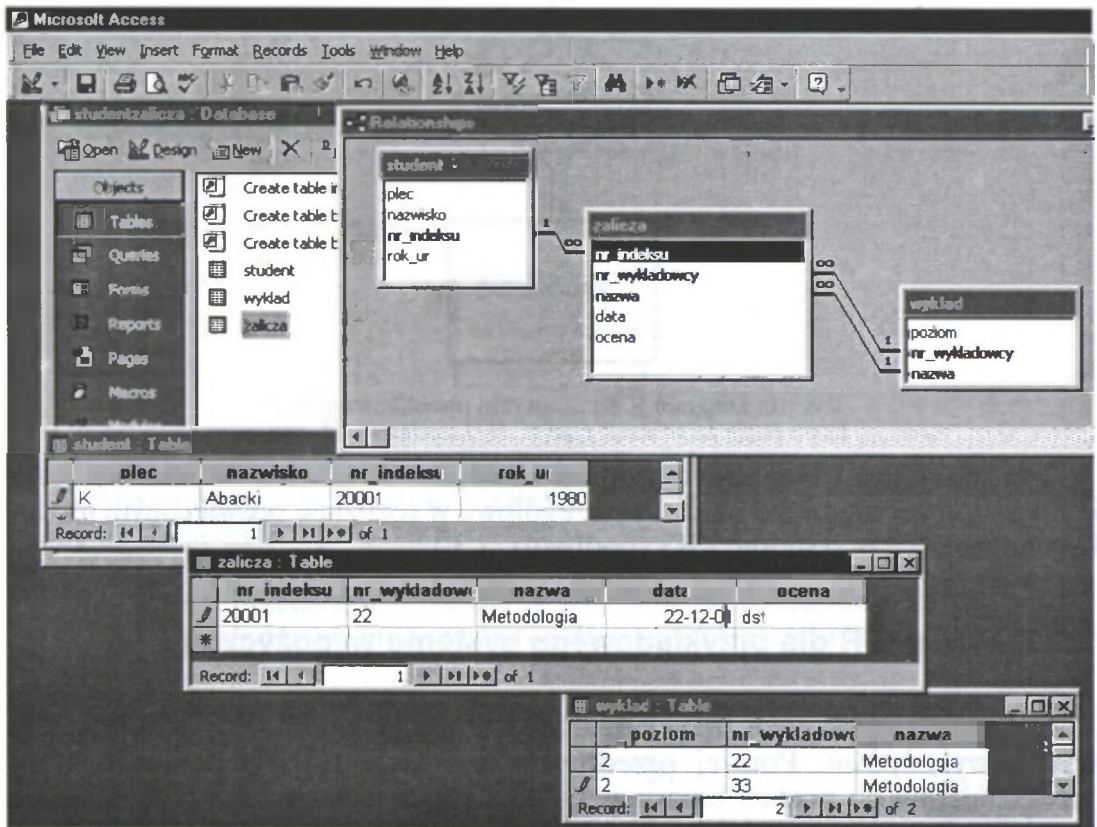
Tabela 2

Struktura danych dla modelu E-R: „STUDENT ZALICZA WYKŁAD”

ENCJA/RELACJA (atrybuty)	Typ danych	Roz- miar	Klucze	Komentarz
STUDENT				
plec	char ¹	2		Płeć studenta
nazwisko	char	30		Nazwisko studenta
nr_indeksu	char	10		Numer albumu
rok_ur	date ²	-		Rok urodzenia
WYKŁAD				
poziom	char	2		Poziom wykładu
nr_wykładowcy	char	10		Identyfikator wykładowcy
nazwa	char	30		Nazwa przedmiotu
ZALICZA				
nr_indeksu	char	10	STUDENT.nr_indeksu	Numer albumu
nr_wykładowcy	char	10	WYKŁAD.nr_wykładowcy	Identyfikator wykładowcy
Nazwa	char	30	WYKŁAD.nazwa	Nazwa przedmiotu
Data	date	n/a		Data zaliczenia
Ocena	char	10		Otrzymana ocena

¹ Pole typu znakowego.
² Pole typu data.

Opisany na rys. 6 i 7 oraz w tabeli 2 schemat można zaimplementować również w typowo relacyjnym systemie zarządzania. Rys. 9 przedstawia implementację naszej przykładowej bazy Student_Zalicza_Wykład w systemie ACCESS firmy Microsoft.



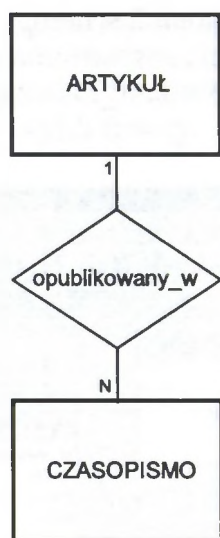
Rys. 9. Implementacja modelu E-R: „Student zalicza wykład” (z przykładowymi danymi) w typowo relacyjnym systemie zarządzania bazą danych Access

Jak widać na rys. 9, w celu stworzenia relacji między encjami STUDENT oraz WYKŁAD, konieczne było stworzenie dodatkowej tabeli.

1.5.2. Bazy bibliograficzne

Kolejną przykładową bazą danych będzie typowa bibliograficzna baza danych. W bazie tej pragniemy zachować informacje o artykułach z czasopism (*analityczny poziom opisu*), opisy bibliograficzne książek (*monograficzny poziom opisu*), i/lub czasopism (*poziom opisu serii*).

Aby poza polami specyficznymi dla artykułu (autor, tytuł, temat itd.) uniknąć przechowywania w każdym rekordzie wszystkich informacji dotyczących czasopisma (takich jak: tytuł czasopisma, numer, wydawca, wolumin, data wydania itd.), łączymy encje typu ARTYKUŁ z encjami typu CZASOPISMO. Dzięki temu unikamy powtarzalności danych, a w przypadku błędów (na przykład w tytule czasopisma) możemy wprowadzić zmiany tylko w jednym rekordzie. Zastosowanie modelu E-R jest dla tego przypadku pokazane na rys. 10.



Rys. 10. Diagram E-R: „...artykuł opublikowany w...”

Typ relacji jest 1:N (jedno czasopismo może zawierać wiele artykułów). Przykładem tym zajmiemy się bardziej szczegółowo w rozdziale poświęconym modelowaniu danych w systemie ISIS (podrozdz. 2.3.).

1.5.3. Model E-R dla przykładowego systemu wypożyczeń

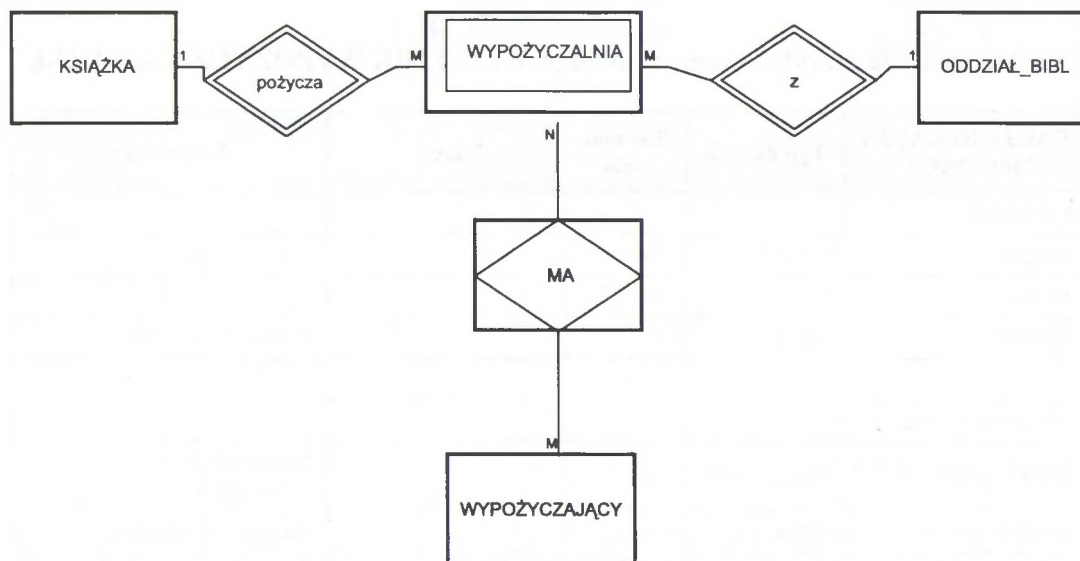
Ostatnim omawianym przykładem zastosowania modelowania E-R będzie system wypożyczeń. Poniżej przedstawiamy relację trójargumentową, którą przekształcimy do postaci binarnej. Zakładamy, że relacje tego typu znacznie lepiej jest reprezentować jako binarne, wykorzystując tzw. słabe encje. Nasz początkowy model wygląda następująco:



Rys. 11. Diagram E-R: relacja trójargumentowa „WYPOŻYCZANIE KSIĄŻEK”

Innymi słowy, jest to baza danych gromadząca dane o czytelnikach wypożyczających książki w różnych oddziałach biblioteki (CZYTELNIK WYPOŻYCZA KSIĄŻKĘ).

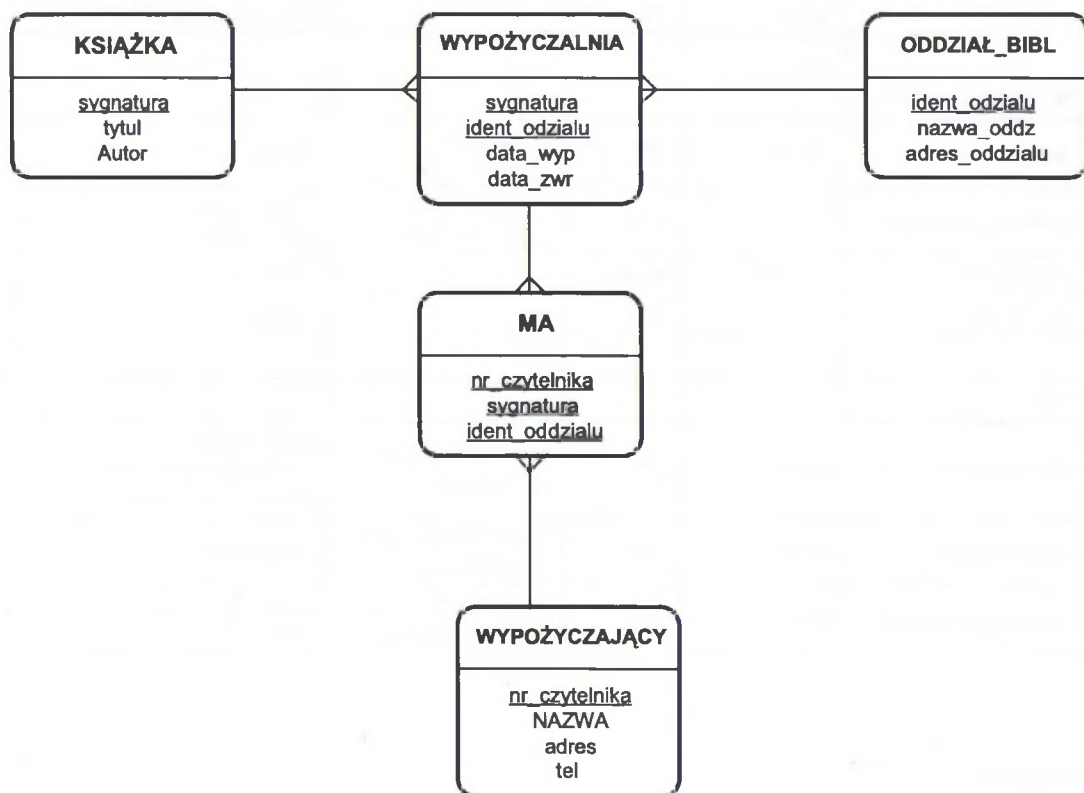
W naszym rysunku WYPOŻYCZA jest relacją trójargumentową (trzech elementów/encji) i w sposób oczywisty winna być sprowadzona do relacji tabeli z kluczem podstawowym, składającym się z trzech encji. Jeśli uczynimy z relacji WYPOŻYCZA słabą encję połączoną relacją z silną encją CZYTELNIK, to uzyskamy taki sam efekt.



Rys. 12. Diagram E-R: relacja binarna ze słabą encją

Warto zauważyć, że związek WYPOŻYCZA został w relacji binarnej zamieniony na słabą encję i połączony relacją *pożyczka* z encją KSIĄŻKA oraz relacją *z* ODDZIAŁ_BIBL.

Dla jasności przedstawmy rozkład własności z podkreśleniem kluczy. Bardziej dokładny rozkład atrybutów (własności) przedstawia tab. 3.



Rys. 13. Diagram E-R: relacja binarna ze słabą encją (z atrybutami i kluczami)

Tabela 3

Układ danych dla przykładowego modelu „CZYTELNIK WYPOŻYCZA KSIĄŻKĘ”

ENCJA/RELACJA (atrybuty)	Typ danych	Rozmiar pola	Klucze	Komentarz
KSIAŻKA				
sygnatura	Char ³	20		Sygnatura
tytuł	Char	30		Tytuł książki
Autor	Char	30		Autor książki
ODDZIAŁ_BIBL.				
ident_oddziału	Smallint ⁴	n/a		Identyfikator oddziału
nazwa_oddz.	Char	15		Nazwa oddziału
adres_oddziału	Char	40		Adres oddziału
WYPOŻYCZAJĄCY				
nr_czytelnika	Char	10		Numer karty czytelnika (karty bibliotecznej)
NAZWA	Char	30		Nazwisko czytelnika
adres	Char	40		Adres czytelnika
tel.	Char	10		Telefon czytelnika
WYPOŻYCZALNIA				
sygnatura	Char	20	KSIAŻKA. sygnatura	Sygnatura
ident_Oddziału	Smallint	n/a	ODDZIAŁ_BIBL. ident_oddziału	Identyfikator oddziału
data_wyp.	Date ⁵	n/a		Data wypożyczenia
data_zwr.	Date	n/a		Data zwrotu
MA				
nr_czytelnika	Char	10	WYPOŻYCZAJĄCY nr_czytelnika	Numer karty czytelnika (karty bibliotecznej)
sygnatura	Char	20	WYPOŻYCZALNIA sygnatura	Sygnatura
ident_oddziału	Smallint	n/a	WYPOŻYCZALNIA ident_oddziału	Identyfikator oddziału

³ Pole typu znakowego.⁴ Pole typu liczbowego.⁵ Pole typu data.

1.6. Ćwiczenia

E-R



Ćwiczenie 1

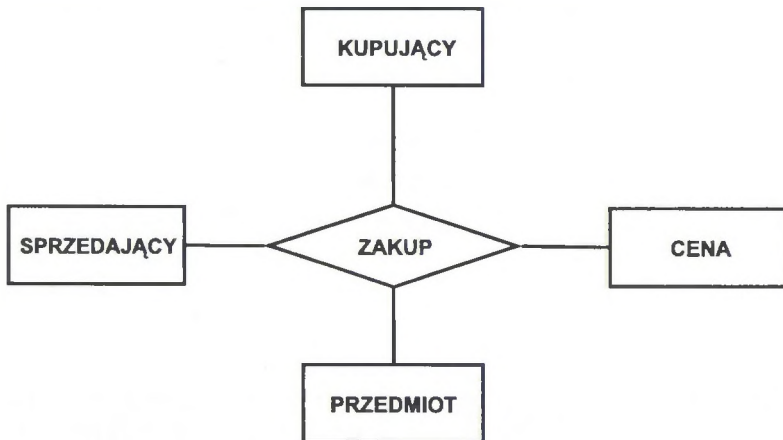
Proszę podać przykładowy układ danych dla przykładu pokazanego na rys. 10. Diagram E-R „...artykuł opublikowany w...”.

Ćwiczenie 2

Zaproponować mechanizm reprezentowania relacji taksonomii (podział, klasyfikacja) w modelu E-R, pokazać jego możliwości oraz podać przykład.

Ćwiczenie 3

Rozważmy diagram przedstawiony na rys. 14. Relacja zakup jest trójargumentowa. Proszę przekształcić ten diagram do postaci równoważnej, w której występują wyłącznie relacje binarne. Czy w tym diagramie można uniknąć redundancji (powtarzania się danych) związanej z obiektami KUPUJĄCY i SPRZEDAJĄCY (osoba kupująca w jednej transakcji może wystąpić jako sprzedająca)?



Rys. 14. Diagram E-R: „Sprzedający kupujący”

2. System CDS/ISIS – wybrane aspekty

2.1. Wstęp

Pragniemy podkreślić, że rozdział ten nie stanowi próby omówienia systemu CDS/ISIS. W przypadku jakichkolwiek wątpliwości zalecane jest zapoznanie się z pełną dokumentacją systemu CDS/ISIS. W celu wygody czytelnika do podręcznika został dołączony aneks zawierający wyciąg z dokumentacji systemu i uwzględniający najistotniejsze aspekty systemu ISIS ze względu na system WWW-ISIS (Aneks B). Nie zwalnia to jednak dociekliwego użytkownika od dokładnego zapoznania się z pełną dokumentacją, która może okazać się niezbędnym źródłem informacji w kwestiach budzących wątpliwości. Zakładamy też, że pod koniec tego rozdziału i zapoznaniu się przynajmniej ze wspomnianym aneksem czytelnik będzie w stanie zaprojektować i wykonać w pełni funkcjonalną bazę danych systemu ISIS.

Należy pamiętać, że w systemie WWW-ISIS system CDS/ISIS (wersja dla systemu DOS, jak i Windows) stanowi tylko element pomocniczy choć niezbędny do administrowania bazami i dotyczy takich funkcji zarządzania bazami danych jak:

- tworzenie i projektowanie bazy danych czyli zbioru głównego i odwróconego;
- okresowa reorganizacja zbioru głównego;
- okresowa aktualizacja zbioru odwróconego;
- eksport oraz import danych;
- złożone raporty.

UWAGA:

W przypadku tworzenia prostych raportów, wprowadzania danych, wydruków, wyszukiwania itd. system CDS/ISIS jest zbędny.



Należy podkreślić, że dokonujemy tutaj bardzo wyraźnego rozróżnienia między systemem zarządzania bazą danych a samą bazą danych. W przypadku udostępniania baz ISIS w sieci Internet (WWW), możliwe jest wykorzystanie w celach administracyjnych zarówno wersji dla systemu operacyjnego DOS, jak i Windows, choć zaleca się używanie pakietu pracującego w systemie DOS (otwieranym jako okno w trybie MS-DOS)¹.

Zależności między WWW-ISIS, CDS/ISIS dla DOS oraz Windows pokazuje tablica 2. Dla uproszczenia zaznaczono wersje: DOS, Windows, WWW-ISIS, WWWISIS (WXIS) oraz JavaISIS i celowo skupiono się na wybranych funkcjach zarządzania bazami danych: wyszukiwaniu, wprowadzaniu danych, imporcie, eksporcie, udostępnianie w sieci WWW i innych narzędziach administracyjnych. Wybrane funkcje poszczególnych pakietów oceniano w pięciostopniowej skali ocen (od 1 do 5). W przypadku pakietów WWWISIS, WXIS wymagane jest dodatkowo tworzenie za każdym razem stosownego oprogramowania. Pamiętajmy, że celem pakietu WWW-ISIS nie jest zastąpienie systemu ISIS wersji dla

¹ We właściwościach programu należy upewnić się, czy są poprawne ustawienia pamięci systemu – dla wszystkich opcji zaleca się wybór automatyczny (więcej patrz aneks B).

DOS lub Windows, a jedynie ich poważne uzupełnienie o zasadnicze funkcje, takie jak wyszukiwanie i wprowadzanie danych przez Internet.

Tablica 2

WWW-ISIS na tle innych systemów obsługujących bazy ISIS

Wybrane funkcje baz danych	DOS ISIS	Win ISIS	WWW-ISIS	WWWISIS WXIS (BIREME)	JavaISIS
 Wyszukiwanie danych	3.5	4	5	wp ²	4
 Wprowadzanie danych	3	3,5	5	wp ³	-
 Wydruki, raporty	4,5	5	4,5 ⁴	3 ⁵	-
 Narzędzia administracyjne	5	4,5	-	wp	-
 Wspomaganie projektowania	3	4	-	4,5	-
 Eksport danych	4	4	4	wp	-
 Import danych	3	3	-	X	-
 Udostępnianie w sieci WWW	-	-	5	5	5
 Wymagania w zakresie dodatkowej wiedzy programistycznej	nie są konieczne Pascal ⁶		żadne	wysokie	-
wp Wymagane programowanie					

² WWWISIS – można przygotować interfejs wyszukiwawczy stosunkowo prosto, przy użyciu systemu GENISIS. WXIS – należy programować w specjalnym ISIS-owym skrypcie. Można niewątpliwie uzyskać bardzo dobre efekty.

³ Bardzo trudne do programowania (autorzy nie znają żadnego dobrze zrealizowanego przykładu).

⁴ Z natury rzeczy raporty muszą być ograniczone co do wielkości.

⁵ Wymaga oprogramowania, z natury rzeczy obowiązują ograniczenia.

⁶ Jest to dodatkowa opcja w przypadku konieczności tworzenia niestandardowych funkcji.

Należy się kilka słów wytłumaczenia odnośnie podstawowych różnic między WWWISIS (BIREME) oraz WWW-ISIS. **WWW-ISIS jest interfejsem systemu zarządzania bazami danych (ISIS) zaś WWWISIS (BIREME) jest sam systemem zarządzania bazami danych.** W rzeczywistości WWW-ISIS może być postrzegany jako możliwy do zmian interfejs.

System CDS/ISIS jest niezwykle popularny w krajach Ameryki Łacińskiej, nieco mniej w Azji i w krajach Afryki (choć również tam jest często jedynym dostępnym narzędziem budowania systemów informacyjnych). W Europie cieszy się dużą popularnością w Hiszpanii i we Włoszech, mniejszą w pozostałych krajach Wspólnoty Europejskiej. W Polsce mimo dużej liczby zarejestrowanych pakietów nie jest obecnie dość często wykorzystywany. Nie sposób jednak przecenić jego olbrzymich możliwości. Niechęć do korzystania z niego wiąże się z ograniczeniami interfejsu użytkownika, które jednak stają się nieistotne w kontekście realizacji interfejsu poprzez WEB. Poniżej pragniemy przypomnieć podstawowe jego zalety.

2.1.1. Ogólny przegląd systemu

ISIS jest uniwersalnym pakietem informacyjno-wyszukiwawczym pozwalającym na przetwarzanie, tworzenie i zarządzanie strukturalnymi, nienumerycznymi bazami danych. Istotnymi cechami systemu ISIS jest jego bogaty język wyszukiwawczy oraz możliwość zastosowania potężnego języka formatowania w celu wyprowadzania danych. Znajomość tej właśnie cechy systemu jest niezbędna do pełnego wykorzystania systemu WWW-ISIS.

Zastosowany sposób przechowywania danych pozwala danym tekstowym na przechowywanie częściowo strukturalizowanych danych w sposób niezwykle ekonomiczny: cały tekst jest przechowywany w polach, zaś pola mogą mieć nieograniczoną liczbę wystąpień, oraz nieokreśloną długość, a w razie konieczności pola mogą być podzielone na podpola. Pamiętajmy, że zaleca się stosowanie systemu ISIS pod warunkiem, że będziemy mieli do czynienia z informacją, której głównym składnikiem jest tekst (alfanumeryczne dane uporządkowane w formie pól i podpól). Stąd zastosowanie systemu ISIS odpowiada wszelkiego rodzaju aplikacjom o charakterze faktograficznym oraz typowym bazom bibliograficznym.

Istnieje wiele wcieleń systemu ISIS dla systemów operacyjnych: MS-DOS, Windows oraz Unix, a jego nieodpłatna dystrybucja jest dokonywana przez akredytowane (przez UNESCO) instytucje prawie we wszystkich krajach świata.

Pamiętajmy, że system ISIS nie jest typowym systemem zarządzania bazami danych, przeznaczonym do celów administracyjnych i przetwarzania danych numerycznych ani też systemem relacyjnym.

Nie należy utożsamiać systemu ISIS z aplikacją biblioteczną, mimo że istnieje wiele aplikacji bibliecznych wykorzystujących narzędzia programistyczne systemu ISIS. Nie można też spodziewać się, że system ISIS umożliwi nam automatyczne lub półautomatyczne stworzenie gotowych aplikacji za pomocą tzw. kreatorów.

2.1.2. Do czego można wykorzystać ISIS?

System ISIS pozwala na utrzymywanie do 16 mln rekordów o łącznej objętości ok. 500MB. Oprogramowanie ISIS można zastosować do przetwarzania danych tekstowych przez zdefiniowanie struktur, które następnie mogą być wykorzystane wielokrotnie dla nieograniczonej liczby baz.

Dane są przechowywane w rekordach, z polami, te zaś mogą ewentualnie zawierać podpola. Całe pola, ich części lub przetworzone zawartości pól oraz wszystkie lub jedynie zaznaczone słowa w polach mogą być zarejestrowane w bardzo wydajnym zbiorze indeksującym, wykorzystywanym do szybkiego odnajdowania rekordów w bazie.

Sposób selekcji poszczególnych danych z pól może być wykorzystywany w celu indeksowania, wyświetlania lub konwersji danych. Rekordy z polami, fragmenty pól lub wynik ich przetworzenia mogą być formatowane w celu ich wyświetlania, wydruku lub przesyłania przez Internet; a dzięki formatowaniu można wyprowadzać dane np. w postaci zwykłych zbiorów tekstowych, zbiorów typu RTF (Rich Text Format) lub zbiorów zapisanych w języku HTML.

Przetwarzanie zawartości pól w rekordach, zarówno, jeśli chodzi o generowanie indeksu (zbioru odwróconego), jak i wyprowadzania danych na ekran, wydruk lub dla programu klienckiego realizowane jest na podstawie specyfikacji wyrażonej w bogatym w swej treści tzw. języku formatowania.

Język wyszukiwawczy w ISIS pozwala m.in. na:

- selekcje pól (dzięki zastosowaniu przedrostków/końcówek);
- układ treści w zbiorze indeksującym dzięki zastosowaniu przedrostków;
- maskowanie prawostronne terminów wyszukiwawczych;
- zastosowanie kombinacji operatorów OR, AND, NOT;
- zastosowanie operatorów sąsiedztwa (położenia) słów (wewnątrz wystąpienia danego pola);
- możliwość wyboru terminów ze zbioru odwróconego w procesie wyszukiwawczym;
- wyszukiwanie pełnotekstowe.

Oprogramowanie ISIS dla Windows jest w stanie odwoływać się także do informacji multimedialnych, nadaje się zatem do implementowania prostych multimedialnych baz danych.

Istotną zaletą oprogramowania ISIS jest jego prostota. Z uwagi na wzrastające zainteresowanie systemem w ostatnich latach daje się zaobserwować rozwój wielorakich narzędzi programistycznych, w tym dla projektantów. Z ważniejszych należy wymienić tu:

1. ISISDLL – narzędzie typu API⁷ dla programistów do tworzenia aplikacji w środowisku Windows.

2. Narzędzia dla programistów do tworzenia aplikacji WEB-owych (WXIS, ISISDBC).

3. WWW-ISIS – omawiane w niniejszej pracy narzędzie dla projektanta do szybkiego tworzenia zaawansowanych aplikacji WEB-owych.

Wraz z tymi narzędziami pojawiły się zaawansowane aplikacje ogólnego przeznaczenia:

⁷ Application Programmer Interface.

- wielorakie interfejsy wyszukiwawcze,
 - niezależne interfejsy wprowadzania danych,
 - częściowo zintegrowane systemy zarządzania bibliotekami,
- oraz gotowe programy użytkowe, głównie narzędzia dla administratorów baz danych i projektantów:

- narzędzia do konwersji danych (np. FANGORN omówiony pobieżnie w Aneksie B, ATTICA, CCF-MARC Converter, DBTOISIS, DB3ISO),
- zróżnicowane narzędzia pomocnicze (np. ISISUT, CISIS Utilities).

Zainteresowanych szczegółami dotyczącymi tych programów odsyłamy do strony systemu ISIS: <http://www.unesco.org/>. My natomiast skoncentrujemy się w dalszej części na podstawowych informacjach dotyczących systemu ISIS.

2.1.3. Obszary zastosowań systemu ISIS, zalecenia i przeciwwskazania

Oprogramowanie ISIS nie jest w stanie efektywnie przetwarzać danych numerycznych. Stąd też aplikacje, w których istotą są zaawansowane obliczenia (jak np. w księgowości lub zarządzaniu papierami wartościowymi itd.), nie powinny być tworzone jako bazy ISIS.

System ISIS nie znajduje też zastosowania w:

- wysoce dynamicznych systemach informacyjnych, w których duża liczba danych zmienia się w krótkim czasie (np. systemy finansowe, giełdowe itd.);
- przetwarzaniu danych, w których elementy układu treści (krój pisma, charakterystyka strony itd.) współistnieją wraz z samymi danymi lub są powiązane z poszczególnymi słowami (np. procesory tekstu lub oprogramowanie do składu tekstu).

Oprogramowanie ISIS znajduje zastosowanie w następujących zastosowaniach informatycznych:

- informacja różnego typu (o różnej strukturze pól) ma być zintegrowana w jedną bazę pozwalając jednocześnie na wyszukiwanie z uwzględnieniem różnorodności zgromadzonych danych; jednocześnie istnieje potrzeba zastosowania niestandardowych zaawansowanych rozwiązań;
- użytkownik (lub organizacja) jest zainteresowany rozwiązaniami typu „zrób to sam”, wymagających więcej niż przeciętnych umiejętności, jednak pozwalających w przyszłości na rozwój aplikacji;
- istnieją poważne ograniczenia finansowe.

Oprogramowanie ISIS nie znajduje jednak zastosowania w następujących sytuacjach:

- czynnik finansowy nie stanowi ograniczenia projektu;
- poszukiwane są gotowe (standardowe) rozwiązania.

2.2. Struktura baz danych i architektura systemu

2.2.1. Bazy CDS/ISIS

Bazę danych systemu ISIS tworzy zbiór danych połączonych ze sobą w odpowiedni sposób. Każda jednostka informacji składa się z elementów zawierających szczegółową charakterystykę poszczególnych encji. Na przykład w bazie bibliograficznej opisującej książki, artykuły i czasopisma, encja składać się będzie z takich elementów, jak: autor (lub autorzy), tytuł, data wydania, miejsce wydania itd.

Podstawowe elementy danych przechowywane są w polach. Każde pole posiada niepowtarzalny znacznik numeryczny. Znacznik ten pełni rolę identyfikatora pola. Znaczniki pól przyporządkowywane są poszczególnym polom w trakcie procesu definiowania struktury bazy. W przypadku, gdy jest to wygodne można w polu wyróżnić podpola.

Pole może być opcjonalne, czyli nie musi występować w każdym rekordzie, może też być obowiązkowe (czyli musi wystąpić w każdym rekordzie). Pole może też być powtarzalne (dany rekord może zawierać więcej niż jedno wystąpienie, np. pole *Autor* może wystąpić w rekordzie kilka razy, wówczas gdy opisywana publikacja ma wielu autorów).

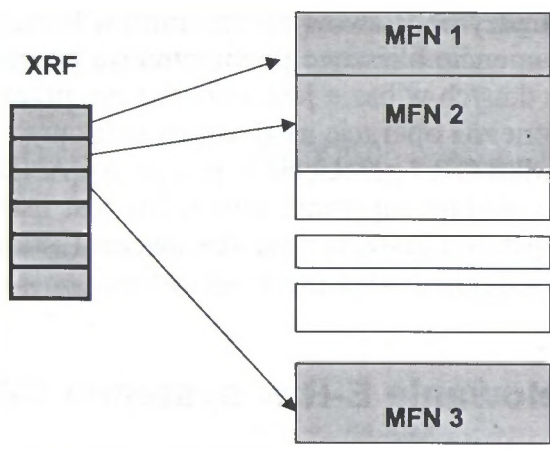
CDS/ISIS został zaprojektowany tak, aby przetwarzać dane w polach o różnej długości. Znaczy to, że jeśli określone pole w danym rekordzie wymaga przechowania tekstu o długości 25 znaków, to długość tego pola będzie wynosiła 25 znaków. Cecha ta pozwala optymalnie wykorzystać pamięć dyskową.

Mimo że CDS/ISIS przedstawia się użytkownikowi jako pojedynczy zbiór informacji, składa się on z szeregu logicznie połączonych i fizycznie odrębnych zbiorów. Zarządzanie tymi plikami jest domeną samego systemu i użytkownik nie musi znać ich szczegółów, aby korzystać z bazy. Jednakże znajomość pewnych faktów dotyczących połączonych zbiorów może lepiej pozwoli zrozumieć funkcjonowanie systemu.

Za definiowanie struktury bazy odpowiedzialna jest usługa ISISDEF. Można też zdefiniować strukturę bazy korzystając ze standardowego edytora tekstowego. Semantyczna struktura bazy danych jest zapisana w Tablicy Definicji Pól (FDT od Field Definition Table). Szczegółowe parametry FDT omówiono w Aneksie B.

Zbiór główny (o rozszerzeniu mst) zawiera wszystkie rekordy danej bazy. Każdy rekord jest identyfikowalny przez niepowtarzalny numer automatycznie przyporządkowany przez system CDS/ISIS w procesie tworzenia rekordu, nazywany Master File Number (MFN). Ponieważ rekordy są zmiennej długości, dotarcie do konkretnego rekordu o ustalonym MFN wymagałoby sekwencyjnego przeglądania zbioru.

Aby umożliwić szybki dostęp do dowolnego rekordu, system CDS/ISIS tworzy specjalny zbiór (o rozszerzeniu xrf) – będący skorowidzem zbioru głównego. Relację między zbiorem głównym (o rozszerzeniu mst) a skorowidzem przedstawia rys. 15.



Rys. 15. Relacje między zbiorem głównym a skorowidzem

Skorowidz jest „wektorem”, w którym i -ta składowa wskazuje adres rekordu $MFN=i$. Ponieważ dostęp do i -tej składowej jest szybki⁸, dostęp do każdego rekordu przez skorowidz jest bardzo szybki.

Aby zapewnić dostęp do rekordów poprzez ich zawartość, CDS/ISIS tworzy tzw. zbiór odwrócony (*inverted file*), będący indeksem całej zawartości bazy. Pozwala to szybko wyszukiwać w bazie rekordy zawierające określone terminy w wybranych polach, np. słowa kluczowe z tytułu, deskryptory, nazwisko autora itp. Schematycznie strukturę zbioru odwróconego przedstawia rys. 16.

Terminy wyszukiwawcze	Numery (MFN) rekordów zawierających terminy
S_1	$Mfn_{11}, Mfn_{12}, \dots Mfn_{1i_1}$
S_2	$Mfn_{21}, Mfn_{22}, \dots Mfn_{2i_2}$
S_3	$Mfn_{31}, Mfn_{32}, \dots Mfn_{3i_3}$
$\dots$	
S_t	$Mfn_{t1}, Mfn_{t2}, \dots Mfn_{ti_t}$

Rys. 16. Relacje między zbiorem głównym a zbiorem odwróconym

Jak widać na rys. 16, zbiór odwrócony zawiera wszystkie terminy, które mogą być wykorzystane jako *punkty dostępu* podczas procesu wyszukiwania, zaś dla każdego z nich jest podana lista rekordów, w których występują te terminy i dodatkowo liczba odwołań do rekordów. Terminy zbioru odwróconego są zorganizowane w formie B-drzewa, co dodatkowo przyspiesza dostęp do poszczególnych słów. Zbiór odwrócony nazywany jest też słownikiem lub skorowidzem.

Wybór terminów do zbioru odwróconego określa się za pomocą tablicy wyboru pól (Field Select Table -FST). W tablicy tej specyfikuje się odpowiednie techniki selekcji oraz polecenia języka formatowania w stosunku do tych pól i ich zawartości ze zbioru głównego, które powinny znaleźć się w zbiorze odwróconym.

⁸ Wszystkie składowe mają tę samą długość na dysku, zatem można policzyć adres i -tej składowej.

Logiczne relacje między podstawowymi zbiorami w bazie CDS/ISIS są bardziej zrozumiałe w momencie bliższego przyjrzenia się procesowi wyszukiwawczemu. Wyszukiwanie danych w bazie jest wykonywane przez dostęp do terminów ze zbioru odwróconego i operacje na zbiorach związanych z tymi terminami (np. scalanie w przypadku OR i przecięcie w przypadku AND). W efekcie, pod koniec całego procesu obsługi zapytania utworzona jest lista rekordów, które spełniają zadane w zapytaniu kryteria wyszukiwawcze. Lista ta jest ostatecznie wykorzystana do wyświetlenia wyszukanych rekordów.

2.3. Modelowanie E-R w systemie CDS/ISIS

Jednym z rozpowszechnionych mitów dotyczących baz CDS-ISIS jest stwierdzenie, że bazy CDS-ISIS są w stanie odzwierciedlić wyłącznie jednorodne zbiory danych, tj. bazy, w których nie ma możliwości pokazywania różnych typów rekordów oraz relacji między nimi. Pokażemy dalej jak można modelować diagramy ER w bazie ISIS⁹.

2.3.1. Typy rekordów (encje)

Jest rzeczą naturalną, że nawet w prostych zastosowaniach pragniemy przechowywać rekordy różnych typów. Na przykład w bazie bibliograficznej¹⁰ możemy przechowywać opisy bibliograficzne książek (monographic level), informacje o artykułach z czasopism (analytical level), i/lub opisy samych czasopism (serial level). Powstaje pytanie, w jaki sposób możemy odróżnić różnego typu rekordy zachowane w bazie ISIS. Otóż rekordy mogą być „rozdzielalne” na różne sposoby, na przykład przez wprowadzenie specjalnego pola, które przechowuje informacje o typie danego rekordu. Z uwagi na funkcje wyszukiwawcze ważne jest, aby rozróżnienie rekordów było odzwierciedlone w strukturze zbioru odwróconego. Możemy do tego celu użyć to pole w ten sposób, że w zbiorze odwróconym wyszczególnione są punkty dostępu do niego, pozwalające szybko wyszukać rekordy określonego typu. Jeżeli w zbiorze odwróconym rodzaj rekordu jest poprzedzony przedrostkiem „TYP=”, to pytania w języku wyszukiwawczym CDS/ISIS będziemy mogli sformułować następująco:

TYP=book * (A=Einstein)

Przykładowo przyjmijmy, że w bazie bibliograficznej pole v11 (o wyróżniku 11) zawiera informacje o rodzaju opisu bibliograficznego (bibliographic level):

AM, AMS, AS – dla rekordów opisujących artykuły z czasopism (AS), lub książek (AM, AMS),

M, MS – dla książek (MONOGRAPH)

S – dla czasopism (SERIAL)

⁹ W aneksie poświęconym systemowi ISIS znacznie bardziej szczegółowo omówione jest zagadnienie języka formatowania, którego znajomość jest niezbędna dla lepszego zrozumienia omawianych tu zagadnień technicznych.

¹⁰ Nie uważamy wcale bazy bibliograficznej za proste zastosowanie.

Chcąc odzwierciedlić tę strukturę w zbiorze FST, możemy zapisać budowanie indeksu dla poszczególnych typów (korzystając z języka formatowania CDS/ISIS)¹¹:

```
1 0 'T=' if V11:'A' then 'ANALYTIC', else if V11.1='M' then  
'MONOGRAPH' else 'SERIAL' fi fi
```

Dzięki temu zabiegowi w przykładowej bazie danych mamy możliwość wyszukiwania określonych typów rekordów. Na przykład, gdy pragniemy wyszukać wyłącznie artykuły autorstwa Smitha, możemy ograniczyć nasze wyszukiwanie przez dodanie T=ANA:

A=Smith * T=ANA

2.3.2. Relacje między rekordami

Załóżmy, że w przykładowej bazie ISIS mamy tylko dwa typy rekordów: „ARTYKUŁ” oraz „CZASOPISMO”. W momencie projektowania struktury dla rekordu „ARTYKUŁ” możliwe są dwa rozwiązania:

1. W każdym rekordzie typu „ARTYKUŁ”, poza polami specyficznymi dla artykułu (autor, tytuł, temat itd.), przechowujemy pełną informację dotyczącą czasopisma (jak: tytuł czasopisma, numer, wydawca, wolumin, data wydania itd.).

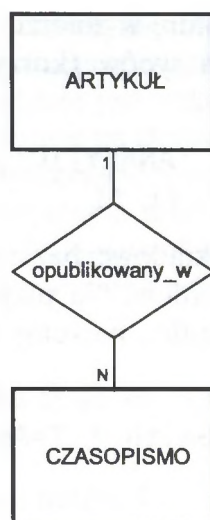
2. W rekordach typu „ARTYKUŁ” przechowujemy wyłącznie elementy poziomu analitycznego tj.: autor, tytuł artykułu, temat, itd.), nie przechowujemy natomiast informacji dotyczących czasopisma, a jedynie połączenie do odpowiedniego rekordu opisującego czasopismo (dokładniej konkretny zeszyt tego czasopisma).

Pierwsze rozwiązanie jest najczęściej stosowanym w przypadku baz ISIS i to dzięki niemu bazy te „zyskują” jednowymiarową strukturę¹². Oczywiście zaletą tego rozwiązania jest możliwość stosowania uproszczonych formatów wyświetlania rekordów wszystkie dane są w bieżącym rekordzie oraz prostsza tablica selekcji danych FST dla zbioru odwróconego. Jednak jego podstawową wadą jest konieczność wielokrotnego wprowadzania tych samych danych do wszystkich rekordów opisujących artykuły z danego zeszytu, a w przypadku regularnego błędu dotyczącego danych związanych z czasopismem istnieje konieczność wprowadzania globalnych zmian.

Podejście drugie unika powtarzalności danych a w przypadku błędów (na przykład w przypadku tytułu czasopisma) możemy wprowadzić zmiany tylko w jednym rekordzie. Przypomnijmy z rozdziału dotyczącego modelowania E-R diagram Diagram E-R „...artykuł opublikowany w...”

¹¹ W systemie WWW-ISIS prefiks „T=” jest zarezerwowany dla odróżniania typów rekordów w procesie wyszukiwawczym.

¹² Warto zauważyć, że z punktu widzenia projektantów relacyjnych baz danych konstrukcja taka nie spełnia wymagań trzeciej postaci normalnej (por. [DAT00]).



Rys. 17. Diagram E-R: „...artykuł opublikowany w...”.

Typ relacji jest 1:N (jedno czasopismo może zawierać wiele artykułów). Połączenia w ISIS-ie mogą być tworzone przy wykorzystaniu numerów rekordów (co nie jest zalecane) lub przez ustawienie specjalnych atrybutów identyfikujących poszczególne rekordy. Atrybut taki może ustalać rzeczywisty rodzaj encji dla każdego rekordu, jak np.: *Numer Identyfikacyjny pracownika*, *Sygnatura książki* itd. lub też może być przyporządkowane automatycznie przez system¹³. Oczywiście zbiór FST jest w tym przypadku znacznie bardziej wyrafinowany. Jeśli pragniemy wyszukiwać artykuły przez tytuł czasopisma, konieczne jest indeksowanie rekordu „ARTYKUŁ” polami znajdującymi się w rekordzie CZASOPISMO. W systemie CDS-ISIS jest to możliwe przez zastosowanie funkcji REF.

2.4. Udostępnianie baz ISIS w sieci WWW

Od samego początku udostępnienia przez UNESCO systemu CDS/ISIS szybko stał się on dużym sukcesem na skalę światową, sprawdzając się w wielu dziedzinach. Głównymi korzyściami płynącymi z wykorzystania aplikacji ISIS były:

- Struktury danych zorientowane na dane tekstowe.
- Uniwersalny język wyszukiwawczy bazujący na technologii zbioru odwróconego.
- Techniki indeksowania, zorientowane na dane tekstowe.
- Język programowania ISIS-PASCAL.
- Język formatowania.
- Łatwość w definiowaniu baz danych.

Wraz z rozwojem Internetu pojawiły się jednak nowe wyzwania i nowe szanse dla aplikacji ISIS. W szczególności powszechne oczekiwania wiązały się z narzędziami umożliwiającymi dostęp do baz systemu CDS-ISIS z Internetu. Pojawiło się wiele rozmaitych prostszych i bardziej zaawansowanych rozwiązań. Jednym

¹³ W systemie WWW-ISIS jest możliwe automatyczne nadawanie pewnych wartości polom, takich jak data, numer rekordu, nazwa użytkownika (na podstawie danych rejestrujących użytkownika itd.).

z pierwszych narzędzi był system Easy-Web. Wkrótce potem, mniej więcej w tym samym czasie, pojawiły się dwa systemy: WWWISIS z BIREME i WWW-ISIS zrealizowany wspólnie przez Instytut Inżynierii Informatycznej i CC (wersja Linuxowa). Na przestrzeni lat 1998-2003 systemy te ewoluowały w różnych kierunkach: (1) na bazie systemu WWWISIS BIREME utworzyła system WXIS, którego głównym elementem jest skryptowy język programowania; (2) we współpracy z FAO rozwinięty został przez Instytut Inżynierii Informatycznej i CC system do szybkiego projektowania gotowych systemów webowych pracujących na bazie systemu CDS-ISIS.

WXIS jest niewątpliwie bardzo interesującym narzędziem przeznaczonym dla programistów. Mimo swych niewątpliwych walorów jego wykorzystanie wymaga wiedzy na temat systemu ISIS, oraz umiejętności programowania, przy czym język programowania jest dosyć specyficzny i nie nawiązuje do znanych języków stosowanych do implementacji zastosowań internetowych, takich jak np. Microsoftowy ASP, czy produkt Cold Fusion.

Inne były cele systemu WWW-ISIS. Autorom systemu WWW-ISIS przyświecały dwa założenia:

1. System powinien być w miarę łatwy dla użytkowników pakietu CDS/ISIS.
2. Użytkownik, nie będąc programistą, z niewielką wiedzą języka HTML oraz bardzo pobieżną znajomością zagadnień sieci WWW, powinien być w stanie bez większych problemów wdrożyć kompletny system informacyjny w sieci WWW działający na bazie danych ISIS.

Temu systemowi jest też poświęcony niniejszy podręcznik. Zakładamy, że na obecnym etapie czytelnik posiada już pewną znajomość baz ISIS (głównie struktury bazy, języka formatowania) oraz jego działania. Warunkiem niezbędnym do właściwego zastosowania systemu WWW-ISIS jest umiejętność stworzenia odpowiedniego zbioru selekcji pól (FST – Field Select Table) oraz odpowiednich formatów wyświetlania danych (patrz rozdz. 3 i 4). Zakładamy, że użytkownik będzie w stanie wykorzystać ważniejsze narzędzia systemu ISIS (zarówno dla wersji DOS, jak i Windows):

- Import/eksport danych (w formacie ISO2709, import danych tekstowych za pomocą takich narzędzi jak FANGORN itp.).
- Tworzenie zbioru odwróconego.
- Uaktualnianie zbioru odwróconego.
- Sporządzanie/odtworzenie kopii zapasowych bazy.

Założenie, że z systemu może korzystać użytkownik z niewielką znajomością technologii WWW wynika z faktu, że dla większości posiadaczy (użytkowników i administratorów) baz ISIS technologia internetowa nie jest najlepiej znana. Z tego też powodu przyjęto, że użytkownik WWW-ISIS nie musi posiadać wiedzy programistycznej: znajomości Perl, cgi, programowania w C, skryptów PHP, ASP, CFL itd. Główny plan działania projektanta tworzącego interfejs webowy dla baz ISIS polega przede wszystkim na definiowaniu szeregu tablic ISIS (podobnych do znanych tablic ISIS – Heurisko). Omawiane w niniejszej książce narzędzie pozwala na stworzenie złożonej aplikacji, składającej się z interfejsu wyszukiwawczego i narzędzi wprowadzania danych bez znajomości języków programowania.

Nie ma konieczności definiowania poszczególnych formularzy wyszukiwawczych w języku HTML, zamiast tego projektant winien zdefiniować swego rodzaju szkielet wspólny dla wszystkich formularzy wyszukiwawczych (wprowa-

dzania danych) w WWW-ISIS, zaś konkretny formularz HTML będzie automatycznie generowany przez system.

Podobnie różne odmiany formularza służącego do wyświetlania danych powinny być zdefiniowane w stosownym „szkielecie”. Reszta wyprowadzanych danych (odpowiedzi ISIS) jest generowana automatycznie przez wykorzystanie języka formatowania ISIS. Atrybuty wyszukiwawcze są definiowane w miarę prosty sposób. Istnieje też łatwy sposób zdefiniowania wielojęzycznego interfejsu (polski, rosyjski, niemiecki, chiński itd.). Bardziej szczegółowe omówienie zasad funkcjonowania systemu przedstawione jest w następnym rozdziale.

2.5. Ćwiczenia

Ćwiczenie 1



Zainstaluj pakiet CDS/ISIS w wersji DOS lub Windows i stwórz prostą bazę danych o następującej strukturze:

CDS/ISIS
FST FDT

Pole Autor – znacznik 10 (pole powtarzalne, składające się z podpól: Imię i Nazwisko)

Pole Tytuł – znacznik 20

Pole Wydawnictwo – znacznik 30

Pole Data Wydania – znacznik 50

Stwórz zbiór odwrócony tak, aby możliwe było przeszukiwanie każdego z wymienionych pól przez odpowiednie przedrostki. Np.: dla pola Autor: „AU=”, dla pola tytuł „TY= ” itd. Przetestuj bazę i wprowadź kilka rekordów. Postaraj się przeprowadzić dowolne wyszukiwanie.



Ćwiczenie 2

Wykorzystując pakiet ISIS (wersja DOS lub Windows) stwórz bazę relacyjną odwzorowującą studentów zaliczających wykłady. Niech baza uwzględnia dwa typy rekordów: „studentów” oraz „wykłady”, zaś w rekordach typu „student” niech nie będą przechowywane informacje dotyczące wykładu, a jedynie połączenie do odpowiedniego rekordu opisującego wykład.

CDS/ISIS
FST FDT



Ćwiczenie 4. Tworzenie FST

Z załączonej bazy Demo (patrz płyta CD) stwórz FST (z odpowiednią techniką selekcji), które pozwoli na pobranie każdego słowa z pola 300 i poprzedzenie go prefiksem A=.

CDS/ISIS
FDT



Ćwiczenie 5. Formaty

Zinterpretuj następujący zapis:

If (p(v400^a) or p(v220) or p(v324) or p(v521))
and (p(v410^a) or p(v540^a) or p(v540^b) or p(v543)) then ' ; ' fi

CDS/ISIS
FORMATY



Ćwiczenie 6. Wyszukiwanie

Załóżmy, że słownik (zbiór odwrócony) zawiera następujące terminy: ROLNICZY BUDŻET, ROLNE PRODUKTY, ROLNICZA STRUKTURA, ROLNICTWO. Znajdź termin, którym można by zastąpić wszystkie wyżej wymienione.

CDS/ISIS
JĘZYK ZAPYTAŃ

Ćwiczenie 7. Wyszukiwanie

Używając podane w poprzednim ćwiczeniu terminy, sformułuj zapytanie, które pozwoli znaleźć wszystkie rekordy dotyczące wszystkiego na temat rynku pracy i rolnictwa w Polsce i na Węgrzech.

CDS/ISIS
JĘZYK ZAPYTAŃ



Ćwiczenie 8. Wyszukiwanie

Sformułuj zapytanie, które znajdzie wszystkie rekordy zawierające dowolną wartość w polu o znaczniku 99.

CDS/ISIS
JĘZYK ZAPYTAŃ



Ćwiczenie 9. Wyszukiwanie

Sformułuj zapytanie, które znajdzie wszystkie rekordy zawierające tekst „rybołówstwo” w polu o znaczniku 6860 lub 6870.

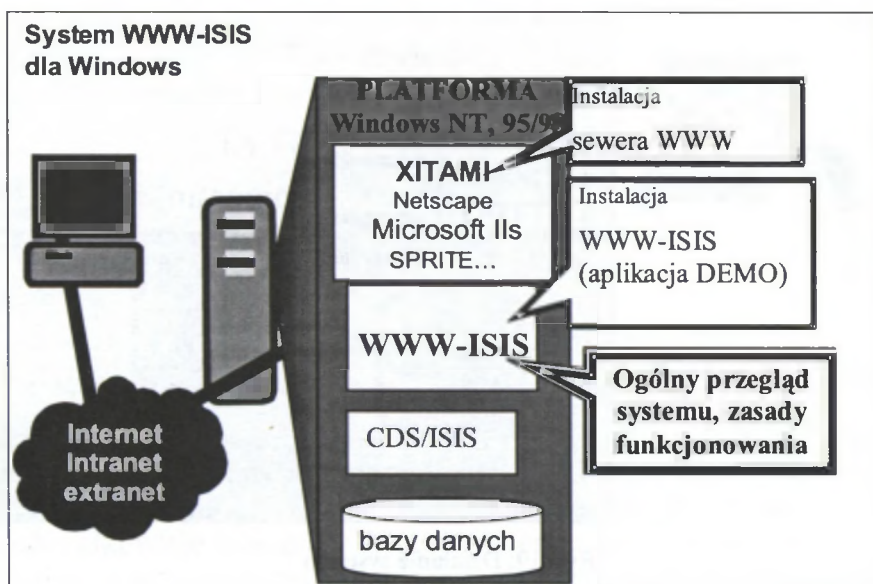
CDS/ISIS
JĘZYK ZAPYTAŃ



3. WWW-ISIS – ogólny przegląd systemu

W rozdziale tym zajmiemy się procedurą instalacyjną samego systemu, omówieniem jego podstawowych funkcji oraz wytłumaczeniem zasad jego funkcjonowania. Wersja demonstracyjna systemu może być wykorzystana jako model wzorcowy jednej z możliwych ścieżek implementacji systemu WWW-ISIS. Do funkcjonowania systemu niezbędne jest uprzednie zainstalowanie serwera WWW (w naszym przykładzie będzie to XITAMI, choć istnieje możliwość instalacji innych serwerów, patrz Aneks E).

Aspekty techniczne są szczegółowo przedstawione w dokumentacji systemu [WWW-ISIS1], zaś oprogramowanie jest dostępne na CD-ROM oraz w sieci Internet. W kolejnych podrozdziałach zostaną przedstawione: szczegóły instalacyjne, podstawowe cechy systemu, jego funkcjonowanie oraz plan działania.



Rys. 18. WWW-ISIS – ogólny przegląd

3.1. Ogólny schemat funkcjonowania systemu

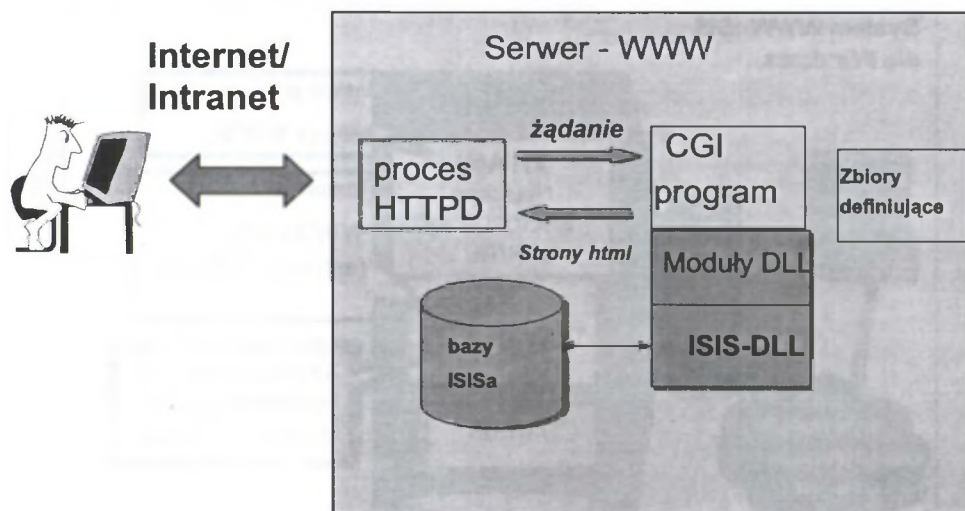
Interakcja systemu podczas formułowania zapytania i wyprowadzania rezultatów wyszukiwań między klientem (przeglądarką WWW), a serwerem WWW wygląda w następujący sposób:

1. Klient aktywuje specjalny proces dzięki programowi cgi, który czyta uprzednio zdefiniowany zbiór, tworzy formularz wyszukiwawczy i przesyła go do klienta.

2. W wyświetlonym formularzu użytkownik wypełnia odpowiednie pola.

3. Użytkownik przesyła wypełniony formularz do serwera.
4. Serwer na nowo uruchamia program cgi.
5. Program cgi przetwarza otrzymany formularz HTML (wraz z wartościami wprowadzonymi przez użytkownika); jeżeli obsługa żądania wymaga dostępu do bazy ISIS, na jego podstawie przesyła odpowiednie polecenia do baz ISIS, program „zwraca” się do modułu ISIS-DLL. Żądanie jest wyrażone w postaci zrozumiałej przez ISIS-DLL – a może nim być: zapytanie ISIS, żądanie przesłania części zbioru odwróconego, żądanie utworzenia nowego rekordu itd.
6. Żądania kierowane przez program cgi wykorzystują moduł ISIS-DLL. W konsekwencji program generuje stronę HTML (formularz zapytania, listę słów kluczowych, komunikaty z bazy danych itd.).
7. Program cgi przesyła dane wyjściowe do serwera WWW w celu ich dalszej interpretacji.
8. Serwer WWW przekazuje dane w postaci zbioru HTML do klienta WWW użytkownika (przeglądarki WWW).

Schematycznie całość jest przedstawiona na rys. 19.



Rys. 19. Działanie systemu

Program cgi jest odpowiedzialny za generowanie i formatowanie danych wyjściowych (na podstawie sformułowanego zapytania), tworzenie zapytania, zarządzanie formularzami itd. Wykorzystuje wiele zbiorów, w których zapisane są istotne informacje o funkcjonowaniu całego systemu.

Przyjrzyjmy się, które zbiory są odpowiedzialne za funkcjonowanie wybranych aspektów systemu.

Zbiory typowe dla baz ISIS (zbiory FST, tzw. formaty wydruku); ich znaczenie w całości funkcjonowania systemu WWW-ISIS jest takie same, jak w przypadku funkcjonowania standardowych aplikacji CDS/ISIS.

Zbiór CFG (isis3w.exe.cfg) zawiera większość istotnych parametrów, ścieżki dostępu, opcje itd.

Zbiory z tablicami definicji wyszukiwań i tablicami przeglądania danych – dla każdego typu rekordu w ISIS-ie, zbiór ten określa odpowiednie elementy w bazie ISIS.

Zbiory HTP – odpowiedzialne za tworzenie szablonów HTML wymaganych przez program cgi w momencie generowania stron HTML; tak więc istnieje search.htp odpowiedzialny za formularze wyszukiwawcze, result.htp odpowiedzialny za tworzenie stron z wynikami itd.

3.2. Serwery i przeglądarki WWW

System WWW-ISIS do prawidłowego funkcjonowania wymaga odpowiedniej platformy serwerowej. Pod tym pojęciem kryją się dwa zagadnienia:

- oprogramowanie do serwera WWW oraz,
- stabilny system operacyjny z rodziny Windows.

Przed podjęciem prób z instalacją systemu WWW-ISIS należy odpowiedzieć sobie następujące pytania:

PYTANIE 1:

Czy system WWW-ISIS będzie pełnił rolę szkoleniową?

☐

TAK

☐

NIE

PYTANIE 2:

Czy będziesz korzystał z samego systemu wyszukiwawczego WWW-ISIS na CD-R?

☐

TAK

☐

NIE

PYTANIE 3:

Czy zamierzamy udostępniać swoje bazy danych w Internecie?

☐

TAK

☐

NIE

PYTANIE 4:

Czy nasz komputer (na którym chcemy udostępniać bazy) jest podłączony do Internetu/Intranetu?

☐

TAK

☐

NIE

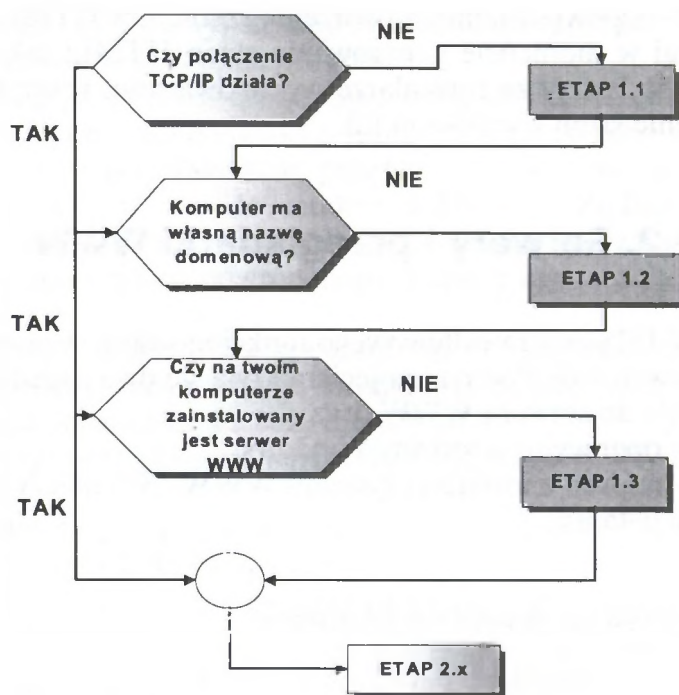
W przypadku ustalenia, że system WWW-ISIS będzie pełnił wyłącznie rolę szkoleniową lub będziemy dokonywali przeszukiwania na CD-R, zagadnienia sieci oraz wspomniana platforma serwerowa stają się drugoplanowe, choć nie można ich całkowicie pominąć. Więcej na temat przygotowywania komputera do pracy w sieci oraz podstawowych spraw dotyczących przetestowanych serwerów WWW, znajdzie czytelnik w Aneksie E.

W dalszej części zakładamy, że bazy danych będą udostępniane w sieci WWW, dlatego przyjęliśmy następujący plan działania – rys. 20.

Wstępny plan implementacji systemu przedstawiony na rys. 20 dotyczy tych aspektów funkcjonowania WWW-ISIS, które związane są bezpośrednio z funkcjonowaniem komputera w sieci Internet/Intranet i nie odnoszą się do implementacji samego systemu WWW-ISIS.

W wielu przypadkach może okazać się, że użytkownik jest podłączony na stałe do Internetu i ma przydzieloną nazwę domenową. W takich sytuacjach można pominąć niniejszy podrozdział i przejść do czynności związanych bezpośrednio z instalacją i tworzeniem interfejsu wyszukiwawczego i/lub wprowadzania danych WWW-ISIS.

Poniższe działania są omówione w dużym skrócie, pełniejsze wyjaśnienie aspektu sieci i serwerów znajdzie czytelnik w Aneksie E.



Rys. 20. Wstępny plan implementacji systemu – sieć Internet/Intranet



ETAP 1.1. PROTOKÓŁ TCP/IP



W systemie operacyjnym Windows NT, połączenie TCP/IP jest jedną z opcji systemu. Sprawdzenie tej opcji jest możliwe z panelu kontrolnego przez ustawienie komponentów sieciowych (konfiguracja). Jeśli opcja ta jest włączona, nie powinno być problemów z połączeniem się z innym komputerem przez sieć TCP/IP, w przypadku wątpliwości radzimy skontaktować się z administratorem sieci.



ETAP 1.2. ADRES URL



W sieci Internet komputer powinien mieć unikalny adres URL. W przypadku wątpliwości w ustaleniu właściwego adresu URL, zaleca się kontakt z administratorem sieci lub z dostawcą usług internetowych.



ETAP 1.3. SERWER WWW



W etapie tym należy upewnić się, czy serwer WWW został zainstalowany na komputerze.

Jeśli użytkownik nie ma doświadczenia z serwerami WWW, zaleca się instalację prostego w obsłudze serwera WWW – XITAMI i przetestowanie najważniejszych jego funkcji.

Po zakończeniu prac nad serwerem WWW i dostępnością w sieci Internet, następuje zasadnicza część pracy nad wdrożeniem systemu WWW-ISIS.

Jako serwer WWW można wykorzystać oprogramowanie NETSCAPE (Netscape Fastrack lub Enterprise WEB server) lub Microsoft WEB serwer IIS. System WWW-ISIS był także przetestowany dla darmowego oprogramowania XITAMI (www.xitami.com). W celu przeszukiwania lokalnych baz danych umieszczonych na CD-R(OM) poleca się darmowe oprogramowanie SPRITE¹.

UWAGA 1:

W celu poprawnego funkcjonowania systemu konieczne jest posiadanie przeglądarki umożliwiającej wyświetlanie:

- tabeli
- ram
- formularzy

Aby w pełni wykorzystać zalety systemu (historia, wyszukiwanie przez tezaurs) przeglądarka winna też umożliwiać wykorzystywanie tzw. cookies.



UWAGA 2:

W celu stworzenia interfejsu wyszukiwawczego dla aplikacji pracujących lokalnie, np. na dyskach CD-ROM, można wykorzystać system operacyjny Windows 95 lub Windows 98. Jednak w celu udostępniania baz w sieci zaleca się instalowanie WWW-ISIS w systemie operacyjnym Windows NT/2000/XP.



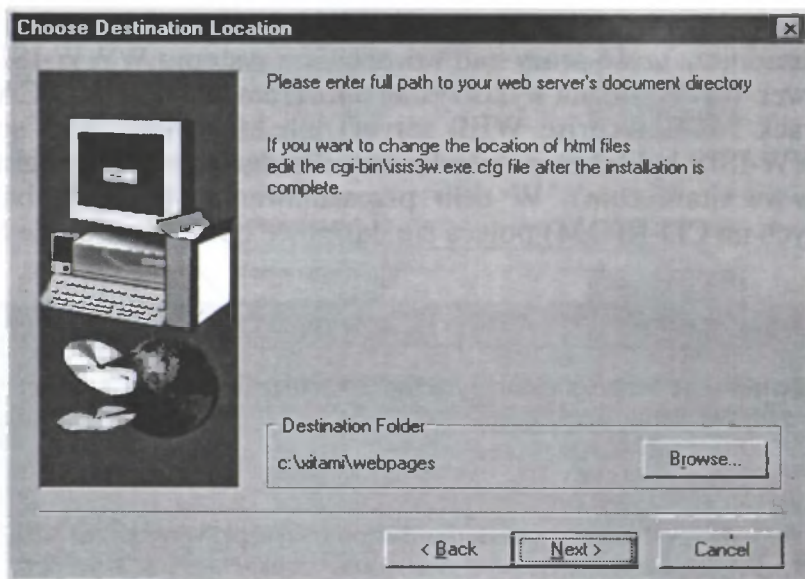
3.3. Instalacja wersji DEMO systemu

Po zainstalowaniu i przetestowaniu serwera WWW, można rozpocząć instalację systemu WWW-ISIS. Sam proces nie powinien nastręczać żadnych trudności, gdyż ogranicza się do uruchomienia programu instalacyjnego umieszczonego na płycie CD-ROM. Należy jedynie pamiętać o określeniu podczas instalacji następujących parametrów:

1. Katalogu macierzystego dokumentów WWW.
2. Katalogu ze skryptami cgi.

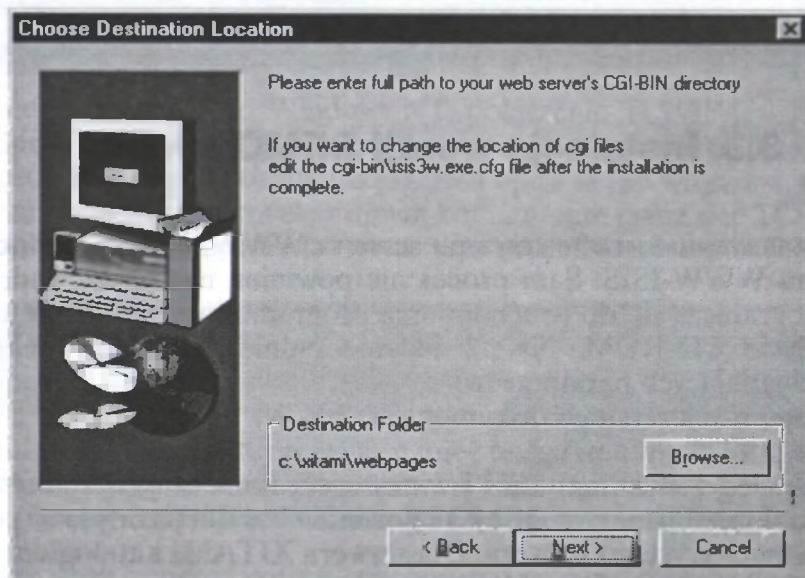
Ad 1. Katalog z dokumentami i większością koniecznych zbiorów WWW-ISIS posiada domyślną nazwę DEMO i powinien być dołączony jako podkatalog do katalogu macierzystego serwera. Dla serwera XITAMI katalogiem macierzystym jest przeważnie C:\Xitami\Webpages (patrz rys. 21), więc system będzie zlokalizowany w C:\xitami\Webpages\Demo. W przypadku IIS katalogiem macierzystym jest przeważnie C:\Inetput\wwwroot\, a więc system będzie zlokalizowany w C:\Inetput\wwwroot\Demo. Jeśli zaś instalujemy system na innym serwerze, należy sprawdzić ustawienia dotyczące jego katalogu macierzystego i zgodnie z nimi dokonywać zmian w czasie instalacji WWW-ISIS.

¹ Więcej na temat środowiska sieciowego oraz wyboru i ustawień serwerów znajdzie czytelnik w Aneksie E.



Rys. 21. Ustalenie parametrów katalogu macierzystego

Ad 2. Kolejny parametr wskazuje na katalog, w którym system przechowuje skrypty cgi. Zakładając, że domyślne ustawienia nie uległy zmianie, w przypadku Xitami jest to katalog C:\xitami\cgi-bin (rys. 22). W przypadku serwera IIS C:\Inetput/cgi-bin konieczne jest uprzednie zdefiniowanie i ustalenie właściwych atrybutów dla katalogu skryptów (por. Aneks E).



Rys. 22. Ustalenie katalogu skryptów cgi

Procedura instalacyjna będzie próbowała rozpoznać nazwę domenową komputera, na którym jest instalowany system WWW-ISIS. Nazwa domenowa (rzeczywista lub fikcyjna) jest istotna, gdyż odgrywa decydującą rolę w zbiorze konfiguracyjnym (isis3w.cfg), a w szczególności jest nieodzowna w przypadku korzystania z opcji systemu odwołujących się do tzw. cookies, pozwalających na korzystanie z takich funkcji systemu, jak historia wyszukiwań („History”), wyszu-

kiwania wykorzystującego tezaurus (*Thesaurus based search*), czy też wprowadzanie danych.

Pod koniec wstępnego procesu instalacji wyświetlane są wszystkie najważniejsze parametry instalowanego systemu. Jeżeli zostaną one zaakceptowane, program instalacyjny rozpoczyna kopiowanie plików instalacyjnych. O postępie procesu instalacyjnego informuje użytkownika okno dialogowe. Wszystkie konieczne zbiory przykładowej bazy DEMO (CDS/ISIS) zostaną skopiowane na komputer użytkownika w taki sposób, aby można było przetestować system.

Po zakończonej instalacji zaleca się ponowne uruchomienie komputera i sprawdzenie funkcjonowania systemu.

Jeśli sieć jest poprawnie skonfigurowana i wszystkie zalecenia instalacyjne zostały wykonane, system powinien działać poprawnie. Wywołanie z okna przeglądarki podanego adresu URL powinno spowodować pojawienie się formularza pokazanego na rys. 29.

Użytkownik demonstracyjnej aplikacji DEMO (zrobionej w systemie WWW-ISIS) powinien mieć do dyspozycji wszystkie podstawowe funkcje systemu, czyli:

- Wyszukiwanie uproszczone (przycisk – Simple search).
- Wyszukiwanie zaawansowane (przycisk – Advanced search).
- Wyszukiwanie za pomocą tezaury (przycisk – Thesaurus based search).
- Historię wyszukiwań (przycisk – History).

W każdym momencie działania systemu użytkownik może korzystać z pomocy systemu i wybrać interesujące go opcje. Na początku można wypróbować wyszukiwanie uproszczone (przycisk – simple search) następnie zaawansowane (przycisk – Advanced search). Zaleca się przetestowanie wyszukiwania za pomocą tezaury, korzystając ze struktury tezaury i wybierając niektóre deskryptory jako terminy wyszukiwawcze.

Jeśli jednak z jakichś powodów instalacja systemu nie zostanie przeprowadzona zgodnie z planem, mogą pojawić się problemy. Poniżej zostaną omówione niektóre z nich.

PROBLEM

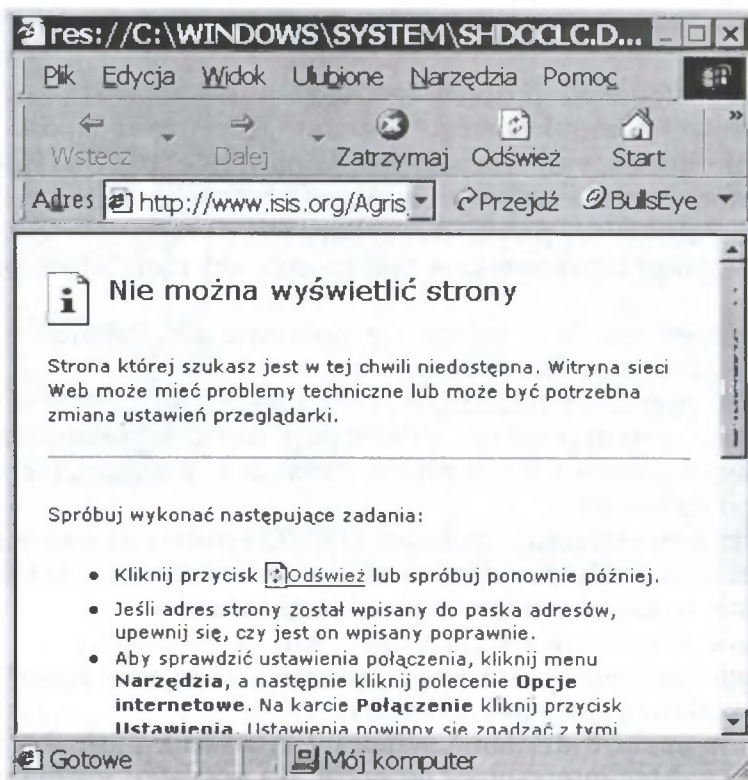
Po wpisaniu odpowiedniego adresu URL w przeglądarce serwer wysyła odpowiedź „Connecting...”, lecz żadna strona nie pojawia się w oknie przeglądarki – „Server could be down” (por. rys. 23).



Należy podjąć następujące działania:

- Sprawdzić adres URL, w szczególności część poświęconą numerowi portu. Przeważnie numer portu wynosi 80, nie mniej należy sprawdzić, czy serwer WWW, na którym jest zainstalowany WWW-ISIS nie korzysta z innego numeru portu.
- Sprawdzić, czy serwer WWW jest uruchomiony. Należy też wtedy sprawdzić konfigurację serwera oraz jego nazwę domenową.

Jeśli działa odwołanie do serwera tylko przez jego numer IP `http://127.0.0.1:<port>/demo`, oznacza to, że komputer nie ma odpowiednich ustawień dotyczących sieci TCP/IP odnośnie nazwy domenowej. W takim przypadku niedostępne będą opcje dotyczące historii wyszukiwań oraz tezaury.



Rys. 23. Problemy z WWW-ISIS – brak dostępu do serwera



PROBLEM

Po wpisaniu adresu URL do przeglądarki WWW system odpowiada „The URL you asked for does not exist on this website” (rys. 24).



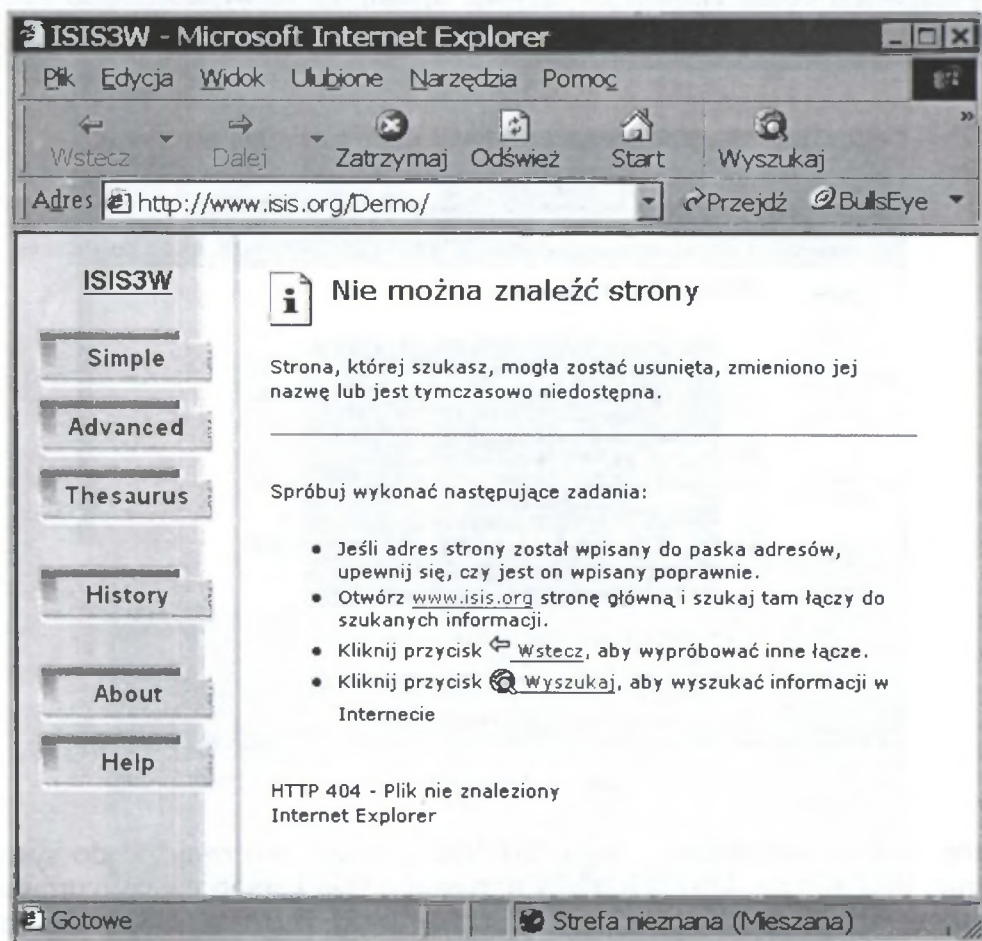
Rys. 24. Problemy z WWW-ISIS – brak zbiorów HTML

Informacja ta oznacza, że sam serwer WWW jest uruchomiony i działa, ale niewłaściwe jest położenie odpowiednich zbiorów HTML na serwerze WWW dla aplikacji DEMO. Prawdopodobnie katalog DEMO nie jest umieszczony w katalogu macierzystym serwera WWW. W takim przypadku radzimy usunąć system WWW-ISIS (przez standardowe opcje w panelu kontrolnym „Dodaj/Usuń”) i ponownie zainstalować system.

Innym możliwym rozwiązaniem tego problemu jest zdefiniowanie dodatkowego katalogu macierzystego na tym samym serwerze WWW. Jednak wymaga to dość dobrej znajomości serwera WWW.

PROBLEM

Po wpisaniu adresu URL do przeglądarki WWW, wyświetlana jest tylko lewa część (lewa ramka), prawa zaś pozostaje pusta (rys. 25).



Rys. 25. Problemy z WWW-ISIS – brak prawej ramki.

Przyczyny takiego stanu rzeczy mogą być dwojakie:

- Katalog cgi dla serwera WWW jest inny niż ten, który został wskazany przez nas w czasie instalacji WWW-ISIS
- Jeśli katalog zawierający skrypty cgi nie jest tym, który został podany w czasie instalacji, serwer WWW nie jest w stanie odnaleźć programu isis3w.exe.

W takim przypadku zalecane jest odnalezienie właściwego katalogu dla skryptów cgi dla naszego serwera i skopiowanie do niego programu isis3w.exe wraz z odpowiednimi zbiorami towarzyszącymi.

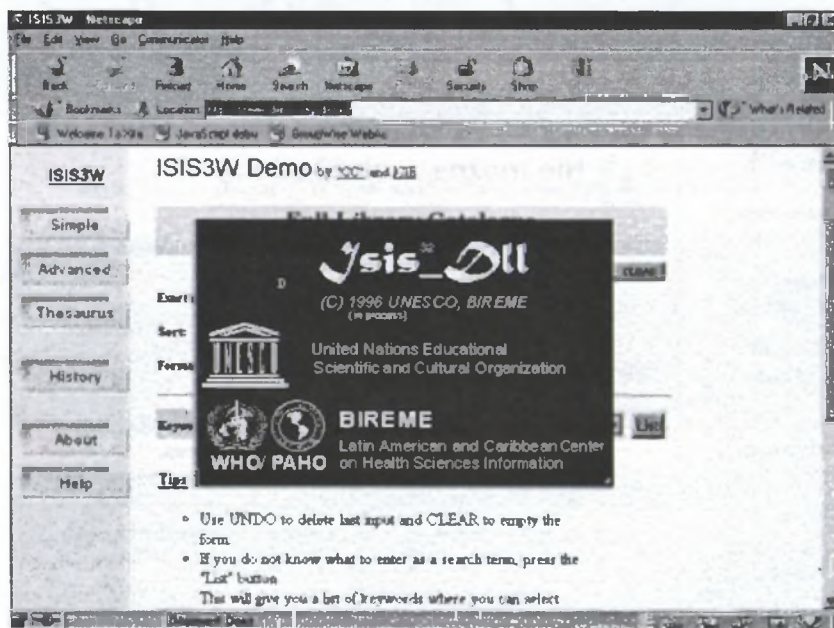
Może się również okazać, że na serwerze WWW nie jest uruchomiona usługa wykonywania skryptów cgi (por. Aneks E). Sytuacja taka nie powinna się zdarzyć w przypadku serwera XITAMI.

Zdarza się, że system działa na lokalnym komputerze (na tym samym, na którym zainstalowana jest przeglądarka oraz serwer WWW), zaś w sieci nie można się z nim połączyć. W takim przypadku należy sprawdzić konfigurację TCP/IP oraz połączenia z siecią (por. Aneks E).



PROBLEM

Ponieważ WWW-ISIS wykorzystuje ISIS-DLL za każdym razem, to może się zdarzyć, że kiedy kierowane będzie zlecenie do serwera, system będzie wyświetlał logo BIREME (rys. 26).

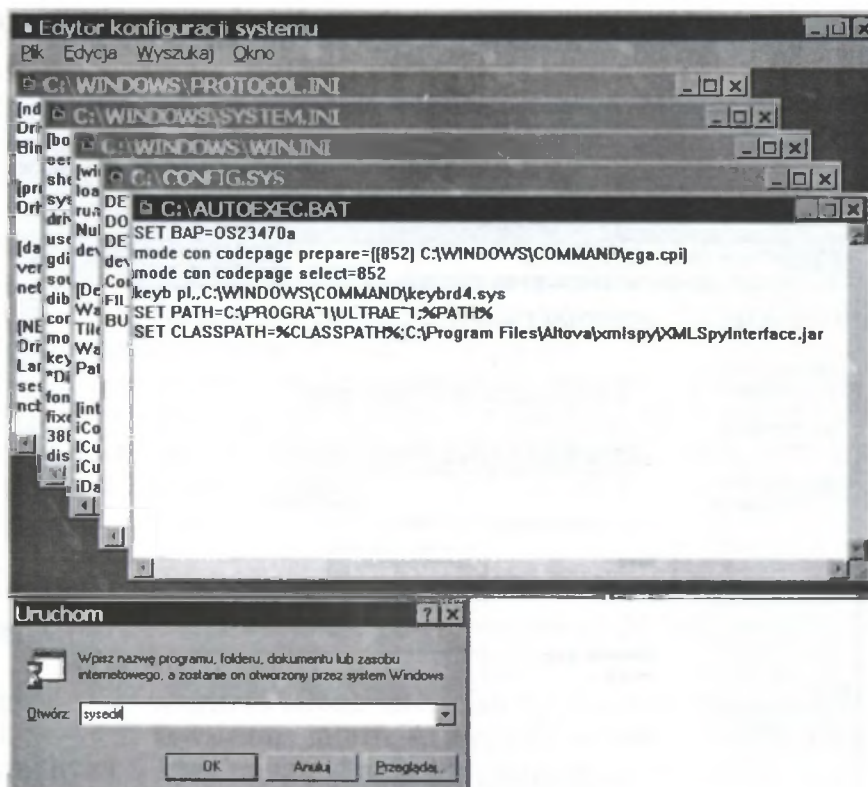


Rys. 26. Logo BIREME

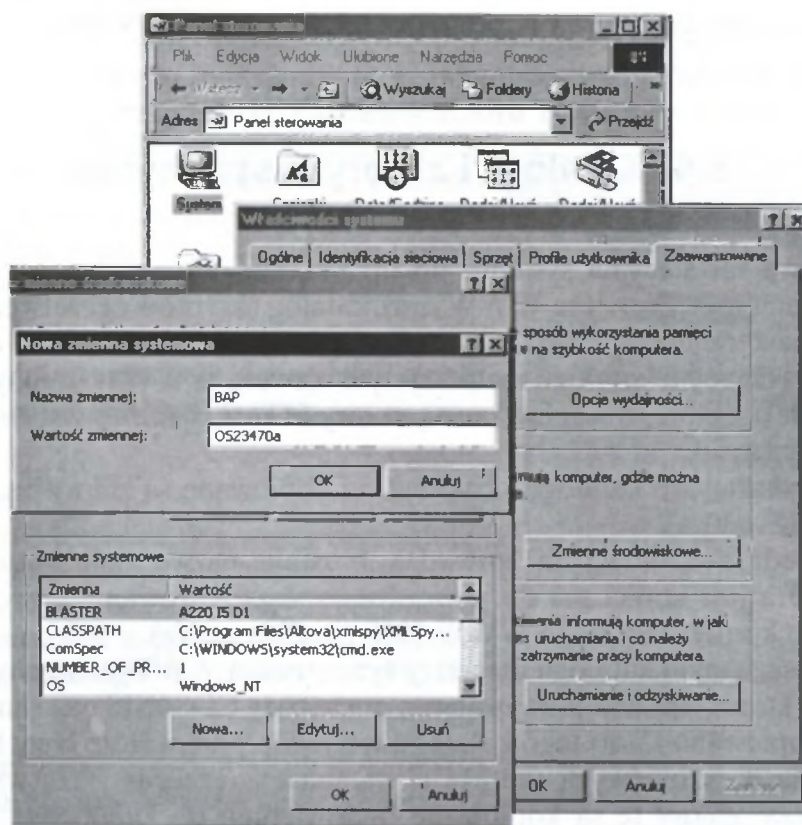
Aby usunąć wyświetlanie logo BIREME, należy wprowadzić do systemu zmienną BAP równą „OS23470a”. W przypadku Windows 9x należy wprowadzić dodatkowy wiersz w pliku AUTOEXEC.BAT : SET BAP= OS23470a (patrz rys. 27).

W systemie Windows 2000 konieczna zmiana ustawień parametrów systemowych odbywa się z poziomu Panelu Kontrolnego (rys. 28). W tym celu należy:

1. Otworzyć panel kontrolny.
2. Wybrać opcję SYSTEM z poziomu Środowisko.
3. Dodać zmienną BAP nadając jej wartość OS23470a.
4. Następnie zaakceptować wprowadzone zmiany przez przycisk OK.

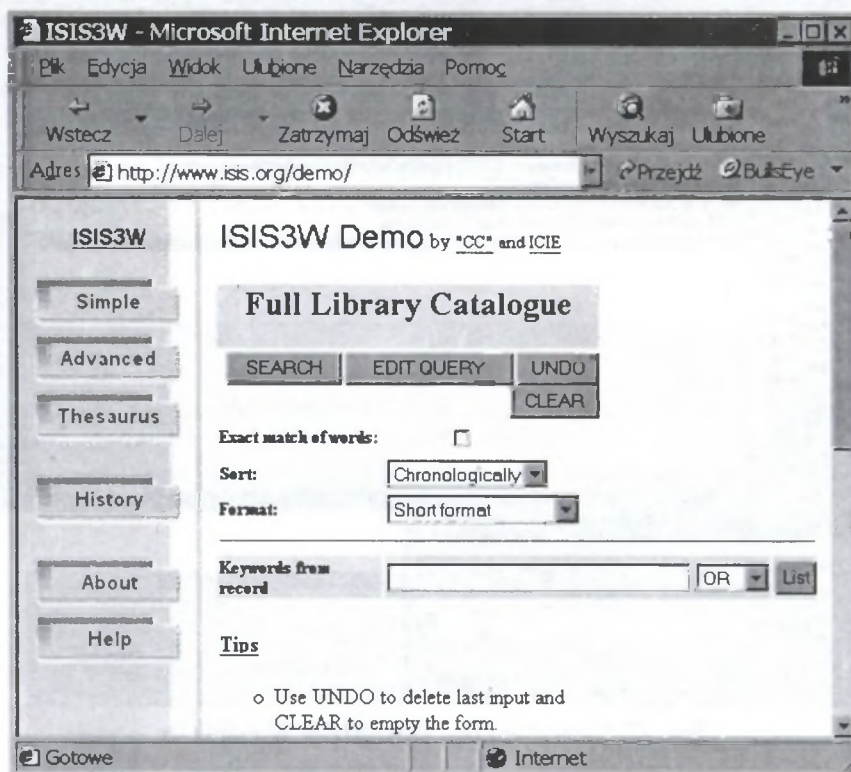


Rys. 27. Usuwanie Logo BIREME w Windows 9x



Rys. 28. Usuwanie logo BIREME w Windows 2000

Po zainstalowaniu systemu przy pierwszej próbie jego uruchomienia powinien pojawić się ekran pokazany na rys. 29.



Rys. 29. Powitalny ekran aplikacji DEMO systemu WWW-ISIS

3.4. Katalogi i zbiory instalacyjne

Procedura instalacyjna systemu po otrzymaniu odpowiednich parametrów (katalog macierzysty serwera WWW oraz katalog skryptów cgi serwera WWW) wykonuje następujące operacje:

1. We wskazanym katalogu macierzystym serwera WWW tworzony jest katalog DEMO, do którego kopiowane są zbiory HTML wraz ze wszystkimi zbiorami definicyjnymi.

2. Do wskazanego katalogu skryptów cgi kopiowane są zbiory `isis3w.exe` oraz `isis32.dll`.

3. Procedura instalacyjna sprawdza i ustala nazwę domenową serwera WWW, która jest konieczna do poprawnego skonfigurowania nowo utworzonego zbioru konfiguracyjnego WWW-ISIS (`isis3w.exe.cfg`).

4. We wskazanym katalogu macierzystym serwera WWW tworzony jest katalog DB, do którego przekopiowywane są puste bazy CDS/ISIS nazwane DEMO. Ponadto kopiowane są do tego katalogu wszystkie pomocnicze bazy ISIS, takie jak: baza haseł wzorcowych (authority database), tezaurus, baza walidacyjna, baza robocza; zbiory te są kopiowane odpowiednio do katalogów: DB\Auth, DB\Thes, DB\Valid and DB\Work.

5. Zbiór konfiguracyjny `isis3w.exe.cfg` jest tworzony i kopiowany do katalogu skryptów serwera WWW.

Ponadto procedura instalacyjna tworzy stosowne skróty na pulpicie serwera (lub w menu startowym). Katalog skrótów zawiera:

- Skrót do interfejsu wprowadzania danych WWW-ISIS.
- Skrót do interfejsu wyszukiwania danych WWW-ISIS.
- Skrót do programu DOS ISIS, który w zamierzeniu przeznaczony jest dla administratora bazy danych (ISIS), dzięki niemu ma on możliwość dokonywania na bieżąco funkcji generowania i uaktualniania zbioru odwróconego, funkcji eksportowania i importowania danych itd.

Przyjrzyjmy się wszystkim komponentom systemu WWW-ISIS, które powinny pojawić się po instalacji systemu. Za przykład niech posłuży struktura katalogów dopasowana do serwera XITAMI. W przypadku innych serwerów struktura powinna być bardzo podobna. Podobnie jak na rys. 30 instalator tworzy katalog Demo z odpowiednimi podkatalogami

Katalog Demo

Katalog Demo zawiera wszystkie potrzebne zbiory HTML oraz zbiory definiujące.

- | | |
|-----------------------|--|
| Index.html | – startowa strona interfejsu wyszukiwawczego WWW-ISIS, |
| Toc.html | – lewa rama interfejsu wyszukiwawczego WWW-ISIS, |
| Kwocinana.html | – lewa rama – strony wyszukiwawcze tezaury, |
| Empty.html | – strona pomocnicza, |
| de.html | – startowa strona interfejsu wprowadzania danych WWW-ISIS, |
| tocde.html | – lewa rama interfejsu wprowadzania danych WWW-ISIS. |

Najważniejszy zbiór to `index.html`, który odwołuje się do `Toc.html` wyświetlanego w lewej ramce przeglądarki. Zawartość zbioru `Toc.html` jest tworzona dynamicznie. Zbiory te będą opisane bardziej szczegółowo w późniejszych rozdziałach. Ponadto folder zawiera zbiory definiujące ekrany wyszukiwania danych oraz ich przeglądania (odpowiednio `database.def` i `browse.def`)

Podkatalog Help

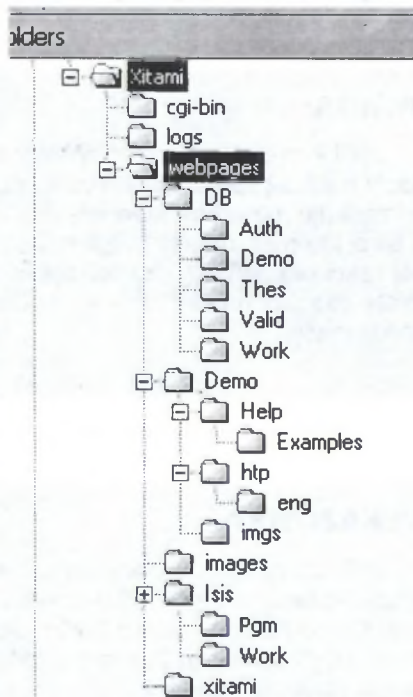
W tym katalogu znajdują się wszelkie zbiory związane z pomocą systemu (w wersji angielskiej).

Podkatalog HTP

Katalog zawiera podkatalog `http` właściwy dla poszczególnych języków. Dla wersji demo dołączona jest tylko wersja angielska „`...\\http\\eng`”. Podkatalog „`...\\eng`” zawiera zbiory HTP, z odpowiednimi wierszami w języku HTML, koniecznymi do dynamicznego tworzenia stron HTML.

Cgi-bin

W czasie instalacji podajemy lokalizację katalogu dla zbiorów cgi dla naszego serwera WWW. Instalator zapisuje w nim następujące programy:



Rys. 30. Aplikacja DEMO – katalogi

- Program cgi isis3w.exe.
- Zbiór isis32.dll, który jest zbiorem typu Isis-dll.
- Zbiór konfiguracyjny systemu będący zbiorem tekstowym, w którym znajdują się podstawowe parametry funkcjonowania systemu, wprowadzone automatycznie podczas instalacji systemu. Nie zaleca się edycji tego zbioru bez znajomości znaczenia poszczególnych parametrów.



UWAGA:

W katalogu webpages (katalogu macierzystym serwera WWW) program instalacyjny automatycznie zainstaluje katalog DB, w którym zostaną umieszczone wszystkie ćwiczebne bazy danych aplikacji DEMO.

DB

W strukturze katalogu macierzystego serwera WWW, procedura instalacyjna tworzy automatycznie katalog DB, w którym są umieszczone wszystkie bazy ISIS aplikacji DEMO:

DEMO – dla bazy DEMO,

AUTH – dla bazy AUTHORITY (jest to baza „wzorów” stosowanych w kontroli jakości danych, takich jak nazwiska autorów, tytuły serii, nazwy autorów zbiorowych, itp.),

THES – dla bazy thesaurusa,

VALID – dla bazy walidacyjnej sprawdzającej poprawność wprowadzanych danych (języki, kraje, kody bibliotek, itp.),

WORK – dla czasowych baz ISIS (konieczny dla przechowywania parametrów podczas sesji pracy z bazą).



UWAGA:

Jeśli komputer (serwer WWW), na którym instalujemy system, nie jest podłączony do sieci i nie ma przydzielonej nazwy domenowej, procedura instalacyjna automatycznie przyporządkuje mu adres www.isis.org. Jeśli w przyszłości zamierzamy uruchomić nasz serwer w sieci Internet, należy pamiętać o zmianie stosownych parametrów w konfiguracji sieciowej serwera WWW, a następnie w zbiorze konfiguracyjnym systemu WWW-ISIS – `isis3w.exe.cfg` należy zmienić ustawienia parametru `CookieDomain` na właściwą nazwę domenową.



WSKAZÓWKA

Jeśli zamierzamy udostępniać nasz interfejs wyszukiwawczy w Internecie a interfejs wprowadzania danych tylko w sieci wewnętrznej (intranet), to zaleca się usunięcie z katalogu Demo zbiorów `de.html` oraz `tocde.html`. Jeśli jednak zamierzamy udostępniać katalog demo użytkownikom Internetu, to należy zadbać o zabezpieczenia dostępu do katalogu Demo (najlepiej hasłem).

3.5. Podstawowe funkcje systemu

System może wykonywać następujące działania:

- Przeglądanie danych (funkcja browse).
- Wyszukiwanie przez zapytanie sformułowane w formularzu HTML.
- Wyszukiwanie przez zapytanie sformułowane w języku wyszukiwawczym ISIS.
- Wyszukiwanie w wielu bazach danych ISIS.
- Zapamiętywanie historii wyszukiwań i ponowne wykorzystywanie wyników wyszukiwań w formułowanych zapytaniach.
- Wyszukiwanie za pomocą tezausa.
- Wprowadzanie nowych danych i aktualizacja danych istniejących w bazie (dostęp do baz autoryzowanych, sprawdzanie poprawności wprowadzanych danych itd.).

3.5.1. Przeglądanie danych (funkcja browse)

Funkcja przeglądania danych (rys. 31) przeznaczona jest dla użytkowników nie mających żadnego doświadczenia z bazami danych oraz niewielkie doświadczenie z Internetem. W przypadku przeglądania danych odnalezienie poszukiwanego rekordu nie jest trudne, o ile znana jest wartość atrybutu, względem którego dane są uporządkowane. Wystarczy jedynie ustalić kryteria uprządkowania wyświetlanych danych oraz ciąg znaków, od którego dane mają być wyświetlane. Na ekranie wyświetlania danych (BROWSE) dostępne są następujące elementy:

- Pole na wpisanie danych, od których ma się rozpocząć przeglądanie (Start from).
- Ustalanie porządku wyświetlania danych.
- Przycisk rozpoczęcia wyświetlania.
- Przycisk przejścia do kolejnej strony.

http://www.isis.org/Demo/

DEMO Thesaurus

Start from: **BROWSE** PgDn

Note: To select a descriptor to your query click on the blue button, to navigate click on the descriptor

- ADENANTHERA
- ADENANTHERA PAVONINA
- AESCHYNOMENE
- AFRICA
- AFRICA SOUTH OF SAHARA
- AFZELIA
- AFZELIA AFRICANA
- AGRICULTURAL DEVELOPMENT

Rys. 31. Opcja przeglądania danych w tezausie

3.5.2. Wyszukiwanie

Funkcja wyszukiwania jest przeznaczona przede wszystkim dla osób mających podstawową wiedzę dotyczącą Internetu oraz baz danych. Poniżej zostaną omówione w skrócie podstawowe cechy systemu.

Wyszukiwanie z podaniem różnorodnych kryteriów wyszukiwawczych oznacza, że użytkownik ma do wyboru w swoim formularzu wyszukiwawczym wiele różnych pól. Do każdego z nich można wprowadzić standardowe operatory ISIS. Pola zaś mogą być połączone operatorami Boolowskimi AND, OR lub NOT (rys. 32 – operatory między polami).

The screenshot shows a search form with the following elements:

- Buttons: SEARCH, SAVE FORM, EDIT QUERY, UNDO.
- Checkbox: Exact match of words: ☒
- Sort: by Author (dropdown)
- Format: Short format (dropdown)
- Keywords from record: [text input] [OR dropdown]
- Keywords from Title: [text input] [OR dropdown]
- Keywords from Serial Title: [text input] [OR dropdown]
- Between the keyword fields are radio buttons for AND, OR, and NOT.

Rys. 32. Formularz wyszukiwawczy: operatory AND, OR łączą terminy wewnątrz pola, zaś operatory AND, OR, NOT łączą pola.

W celu uczynienia procesu wyszukiwawczego bardziej przyjaznym, do każdego z pól dołączony jest swojego rodzaju słownik (rys. 33). Słowniki te mogą być wyświetlane w postaci pól typu „kombi” („combo”) lub tworzone automatycznie na podstawie zbioru odwróconego.

The screenshot shows a keyword selection interface with the following elements:

- Keywords from record: [text input]
- Buttons: Select, PgDn, Home, UNDO.
- List of terms with checkboxes and counts:

☐	ABA	3
☐	ABATTAGE	2
☐	ACADEMIE	2
☐	AGENTS	2
☐	AGGRAVER	1
☐	ALIMENTAIRE	2
☐	ALIMENTARIUS	1
☐	ANALYSE	1
- Buttons: OR, List (multiple instances).
- Radio buttons for AND, OR, NOT.
- ther: [text input] [OR dropdown] [List button]

Rys. 33. Korzystanie ze słownika (przycisk – List)

Wszystkie wybrane terminy ze słownika są automatycznie wprowadzane do zapytania wyszukiwawczego w formularzu.

Rys. 34. Redycja zapytania w języku wyszukiwawczym ISIS (EDIT QUERY)

Dzięki wykorzystaniu funkcji *ISIS Query Language* lub *EDIT QUERY* (rys. 34) stworzone przez wypełnienie formularza zapytanie może być następnie wyświetlone jako zapytanie w języku wyszukiwawczym ISIS i poddane bezpośrednio edycji przez zaawansowanych użytkowników (rys. 34).

W przypadku wykorzystania funkcji *multi-base search* pojedyncze zapytanie może być skierowane do wielu logicznie kompatybilnych baz danych. Dla użytkownika zaś oraz dla operacji indeksowania, sortowania i wyświetlania danych taki zbiór baz danych może być postrzegany jako jedna logiczna baza danych.

Przycisk *History* pozwala na odwoływanie się do poprzednich zapytań dokonanych podczas danej sesji (np. „#2 ^ #1”, rys. 35).

Dla wyświetlanych wyników użytkownik może wybrać stosowny format wyświetlania i ich sortowania.

Dane wyświetlane są na kolejnych stronach.

Rys. 35. Historia wyszukiwań (Opcja „History”)

Interfejs wyszukiwawczy jest bardzo elastyczny i może zaspokoić najbardziej wymagających użytkowników:

- Możliwe jest stworzenie i korzystanie z niezależnych wielojęzycznych interfejsów.
- Projektant może przewidzieć wiele formatów wyświetlania danych i ich sortowania.
- Ekrany (formularze) wyszukiwania są kontrolowane przez odpowiednie tablice definiujące interfejs wyszukiwawczy.
- Wszelkie konieczne definicje wyszukiwawcze mogą być wykonane bez dogłębnej znajomości języka HTML.

3.5.3. Wyszukiwanie za pomocą tezaurya

Ta funkcja okazuje się pożyteczną dla zaawansowanych użytkowników. Dostarcza ona dodatkowych informacji na temat logicznych związków pomiędzy słowami kluczowymi, służącymi do indeksowania i klasyfikacji dokumentów. System pozwala nawigować między różnymi terminami uwzględniając relacje semantyczne między nimi (takie, jak deskryptor węższy, deskryptor szerszy, kojarzeniowy itd.).

AGRIS	Selected Descriptors:
■ ABANDONED LAND - UF: DERELICT LAND - IDLE LAND - BT: ■ <u>LAND</u> - RESOURCES - RT: ■ <u>FARM</u> - CLOSURES ■ <u>LAND</u> - DIVERSION ■ <u>LAND USE</u> ■ <u>MARGINAL</u> - LAND ■ <u>SHIFTING</u> - CULTIVATION ■ <u>WASTE LAND</u>	Environmental diseases ; Abiotic factors ; Abnormalities ; LAND USE SEARCH EDIT QUERY CLEAR

Rys. 36. Wyszukiwanie za pomocą tezaurya

Użytkownik poruszając się po strukturze tezaurya wybiera terminy stosownie do konstruowanego przez siebie zapytania. System jest w pełni zautomatyzowany, zatem w procesie nawigacji możliwe jest wskazywanie terminów (deskryptorów) najbardziej odpowiadających naszemu zapytaniu. Terminy te są zbierane do obszaru tekstowego (rys. 36, prawa strona) w celu wysłania zapytania lub jego reedycji w formularzu lub w samym języku wyszukiwawczym ISIS. Opcja wyszukiwania przez zastosowanie tezaurya daje duże możliwości poprawienia relewancji wyników wyszukiwawczych. Podstawowymi cechami tego modułu wyszukiwawczego są:

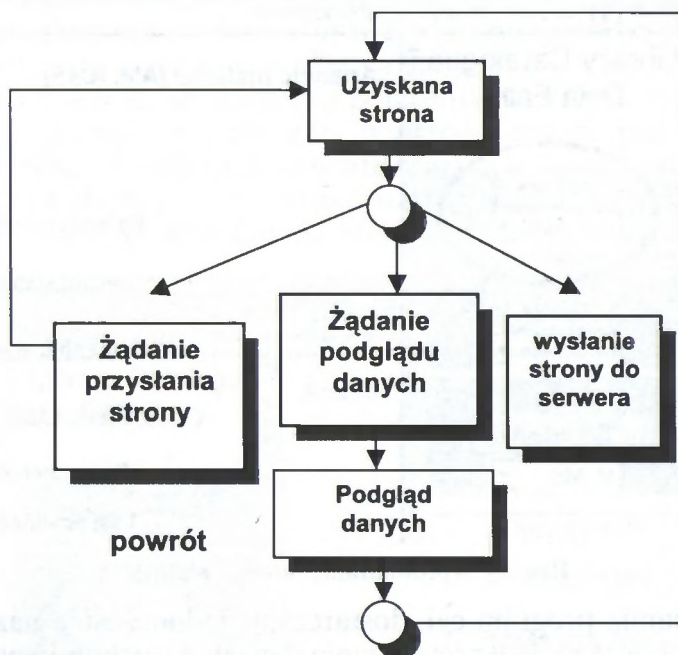
- Różne strategie nawigacji – wybór deskryptorów z listy alfabetycznej, listy KWOC² lub układu systematycznego.
- Nieograniczona liczba tezaurusów współpracujących z nieograniczoną liczbą baz danych oraz typów rekordów.
- Jeśli tezaurus posiada połączenia do innych tezaurusów lub tabel zbieżności, możliwe jest zaimplementowanie jednego systemu tezaurusowego dla wielu baz danych.

Możliwe jest także zaimplementowanie tezaurusów wielojęzycznych w celu wyszukiwania w języku (etnicznym), który jest różny od języka, w którym baza została zaindeksowana. Oznacza to, że możliwe jest przeprowadzenie wyszukiwania w bazie zaindeksowanej deskryptorami angielskimi udostępniając francuskojęzycznemu użytkownikowi francuskojęzyczną wersję tezaurusu.

3.6. Interfejs wprowadzania danych

System WWW-ISIS dostarcza również możliwości zaimplementowania zaawansowanego podsystemu wprowadzania danych. W procesie wprowadzania danych wykorzystuje się formaty ISIS w celu:

- kontroli poprawności na poziomie rekordu (globalna) i pola (lokalna);
- sprawdzenia poprawności usuwanych danych;
- wyboru stosownego formularza;
- sprawdzenia uprawnień do dokonywania zmian w bazie;
- kopiowania wartości domyślnych;



Rys. 37. Schemat interakcji w interfejsie wprowadzania danych

² Keyword Out of Context – wybór teminu z tej listy jest szczególnie wygodny, gdy poszukiwany termin składa się z kilku słów i znane jest słowo, ale nie mamy pewności, że znajduje się ono na początku deskryptora.

Ponadto system wprowadzania danych posiada dodatkowe cechy:

- Dostęp do wybranych funkcji systemu chroniony jest hasłem.
- Można zdefiniować dowolnie wiele formularzy, które mogą być dostosowane do dowolnie wybranych encji, formularze mogą być wielo- lub jednostronicowe.
- Tworzone są dynamicznie tablice podglądowe na podstawie istniejących haseł wzorcowych z wybranych pól w bazach ISIS (*authority files*). Tablice wzorcowe mogą być wybrane z głównej bazy danych lub dowolnej innej, np. z tezaurusa, bazy czasopism, wydawnictw itd. Dzięki przyciskowi LIST możliwe jest wyświetlenie i wybranie stosownych wartości i wprowadzenie ich do pól (patrz rys. 33).
- Można zdefiniować statyczne pola wyboru (*combo*), pola zaznaczania danych (*check*) w celu wybierania wartości z niewielkiego podzbioru danych i ich wprowadzania do podpól. W formularzach można definiować grupy powtarzalne dla pól powtarzalnych ISIS zawierających podpola. System automatycznie tworzy jedno wystąpienie pola dla każdej powtarzalnej grupy (wprowadzając wszystkie konieczne znaczniki podpól).

Formularze są definiowalne dzięki zastosowaniu prostego skryptu o strukturze zbioru XML – zbioru o nazwie entry.def.

3.6.1. Wprowadzanie rekordu

Wprowadzenie rekordu uzyskuje się dzięki przyciśnięciu przycisku „ADD new” (rys. 38). Schemat interakcji między klientem (przeglądarką WWW) a serwerem WWW w interfejsie wprowadzania danych WWW-ISIS wygląda następująco (patrz rys. 37).

The screenshot shows a web browser window with the address bar displaying 'http://www.isis.org.eng:90/libcat/de.html'. The page content is divided into two main panels. The left panel, titled 'Library Catalogue Data Entry', contains a form with several elements: a button labeled 'Add new' (which is circled in red), a dropdown menu currently showing 'AM, AMS', a button labeled 'Update', a dropdown menu labeled 'by browse', a button labeled 'Set defaults', and another dropdown menu currently showing 'M, MS'. Below this panel is a 'LOGIN' button. The right panel, titled 'Analytic material (AM, AMS)', contains a form with several fields: 'Record identifier', 'Record status', 'Bibliographic level', 'Electronic address', 'Date entered', and 'Last modified'. Each of these fields has a small icon to its right, likely representing a help or information button.

Rys. 38. Wprowadzanie nowego rekordu

Klient uruchamia program cgi, dostarczając jednocześnie nazwy odpowiedniego arkusza roboczego do wprowadzania danych. Uruchomiony przez nas program cgi czyta zdefiniowane uprzednio parametry w zbiorze definicji wprowadzania danych i znajduje szczegółowy opis wskazanego arkusza roboczego, wraz z tym opisem tworzy stronę zgodną z podanym arkuszem wprowadzania danych (czyli konkretnym formularzem HTML), a następnie przesyła go do klienta wraz z ustawionymi wartościami domyślnymi w wybranych polach (o ile takie są).

Użytkownik wypełnia w formularzu odpowiednie pola. Jednak przed wysłaniem do serwera rekordu (dane umieszczone są oczywiście w formularzu), można wysłać do serwera jeszcze inne zlecenia:

1. Jeśli dla danego pola istnieje potrzeba przejrzenia już wprowadzonych danych, to uruchamiany jest program cgi wraz z parametrami określającymi, jaka część zbioru odwróconego bazy ISIS powinna być wyświetlona (wykorzystana jest funkcja BROWSE), a następnie program przesyła do użytkownika stronę z danymi do przejrzenia wraz z możliwością zaznaczenia w niej wartości, które mają być wprowadzone do danego pola w formularzu.

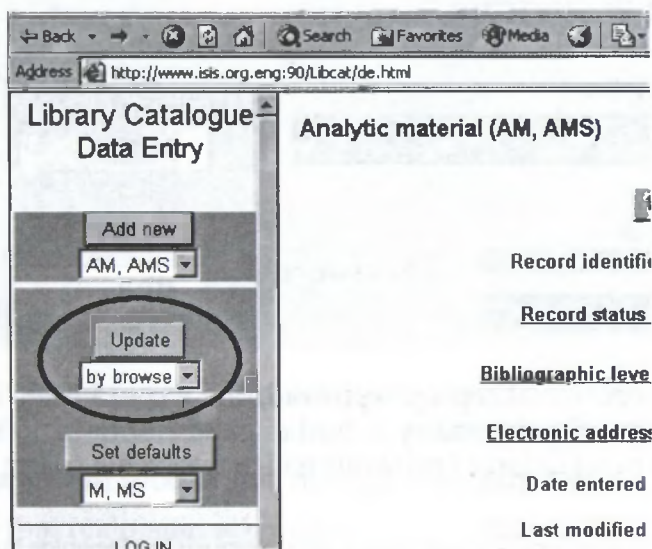
2. Jeśli istnieje potrzeba wyświetlenia innej strony, program cgi jest uruchamiany wraz z parametrem określającym, która strona winna być wyświetlona. W tabeli definicji wprowadzania danych generowana jest stosowna strona z wprowadzonymi danymi (lub z danymi domyślnymi, jeśli strona jest wywoływana po raz pierwszy), następnie system przesyła stronę do klienta.

Jeśli użytkownik prześle wypełniony arkusz (przeważnie składa się z kilku stron) do serwera, serwer WWW ponownie uruchamia program cgi, który tym razem tłumaczy formularz do postaci rekordu w bazie ISIS, a następnie uruchamia format sprawdzania poprawności danych i jeśli nie zostaną wykryte żadne błędy, program zachowuje rekord; jeśli system odnajdzie jakiś błąd we wprowadzonych danych, do użytkownika przesyłana jest strona z informacją o błędzie.

3.6.2. Uaktualnianie rekordu

W przypadku uaktualniania rekordu schemat interakcji jest podobny, jak w przypadku wprowadzania danych (por. rys. 37) z jedną tylko różnicą: zanim pobrana do aktualizacji strona zostanie wysłana do klienta, program cgi czyta wskazany rekord i dokonuje wyboru wartości z poszczególnych pól wprowadzając je do wskazanego arkusza wprowadzania danych.

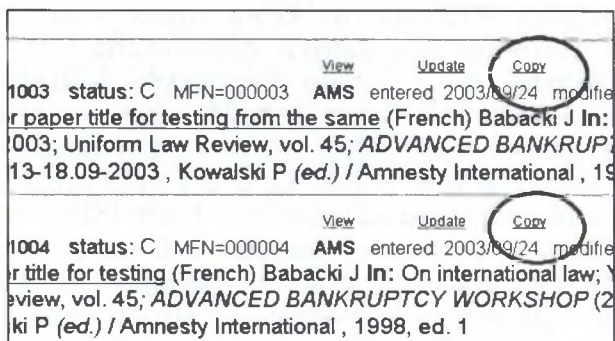
Kolejna różnica polega na działaniu polecenia SAVE (zachowaj). Po udanym sprawdzeniu poprawności rekordu, poprzedni rekord jest modyfikowany zgodnie z zawartością arkusza (pola zawarte w arkuszu wprowadzania danych są wymieniane). Jeśli wykryty jest jakiś błąd w uaktualnianym rekordzie, rekord nie jest zapisywany w bazie, tj. aktualna jest stara wartość rekordu.



Rys. 39. Uaktualnianie rekordu

3.6.3. Kopiowanie rekordu

Schemat interakcji jest połączeniem funkcji wprowadzania nowych rekordów i uaktualniania starych. Jeśli użytkownik wskazuje rekord do skopiowania, zanim strona zostanie przesłana z powrotem do klienta, program cgi czyta dostarczony rekord i tłumaczy go na arkusz wprowadzania danych (podobnie jak w przypadku UPDATE – uaktualniania danych); gdy arkusz jest wysyłany do serwera WWW system próbuje stworzyć nowy rekord (podobnie jak w przypadku opcji ADD NEW – tworzenie nowego rekordu).



Rys. 40. Kopiowanie rekordu

3.6.4. Testowanie systemu wprowadzania danych

Po zainstalowaniu systemu należy go przetestować. Poniżej przedstawione zostaną wybrane możliwości systemu wprowadzania danych WWW-ISIS z punktu widzenia końcowego użytkownika systemu. System jest uruchamiany wraz z podaniem w panelu przeglądarki adresu <http://www.isis.org/Demo/de.html>³.

System powinien uruchomić się wraz z wyświetleniem strony rejestracyjnej (rys. 42). Zaraz po zainstalowaniu domyślne ustawienia wprowadzania danych są następujące:



REJESTRACJA W SYSTEMIE

Nazwa użytkownika:

User-ID =

dba

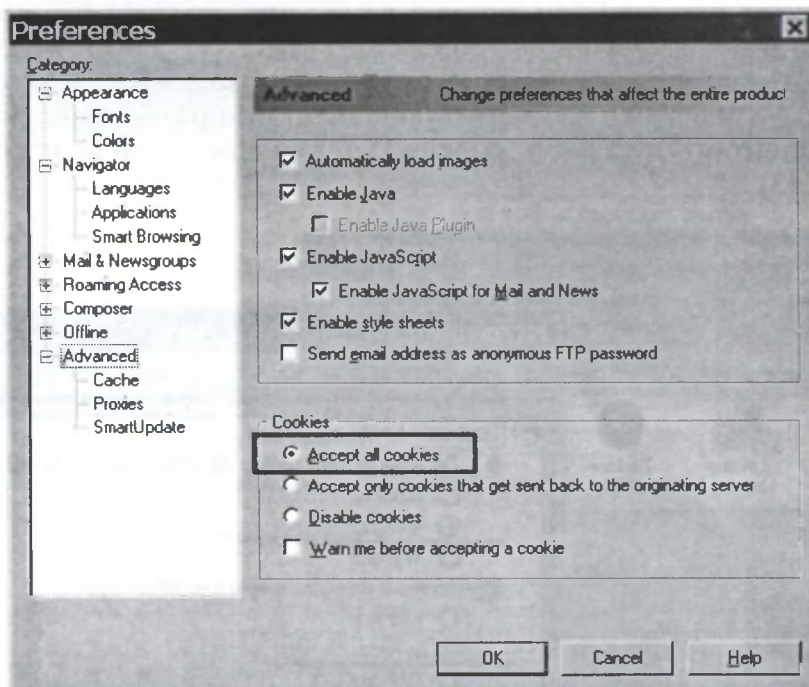
Hasło użytkownika:

Password =

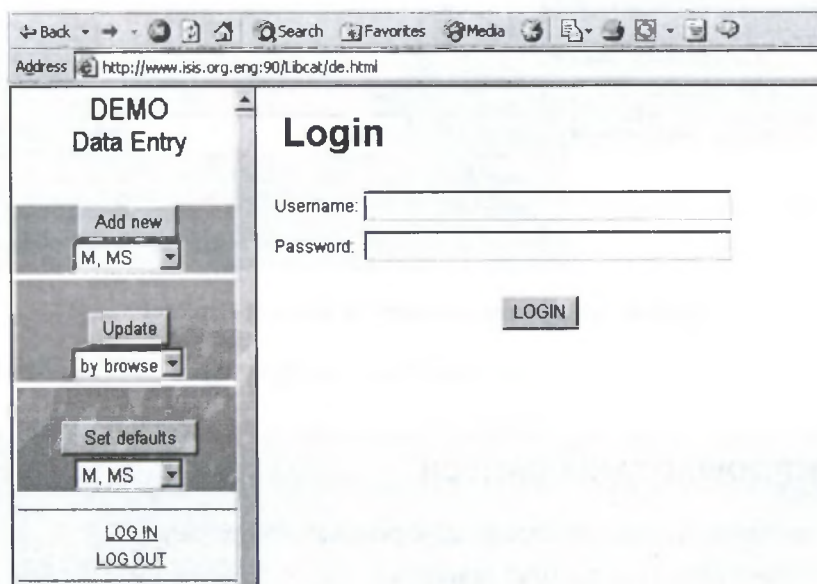
dba

Jeśli system pozornie akceptuje wprowadzoną nazwę i hasło i następnie żąda ponownego wprowadzenia nazwy i hasła, należy sprawdzić ustawienia tzw. cookies w naszej przeglądarce i pozwolić na ich tworzenie (patrz rys. 41 oraz 43).

³ Uruchamiamy system wpisując jego nazwę <http://www.isis.org/Demo/de.html>, jeśli został zainstalowany na komputerze bez ustalonej nazwy domenowej.



Rys. 41. Ustawienia „cookies” w Netscape Navigator

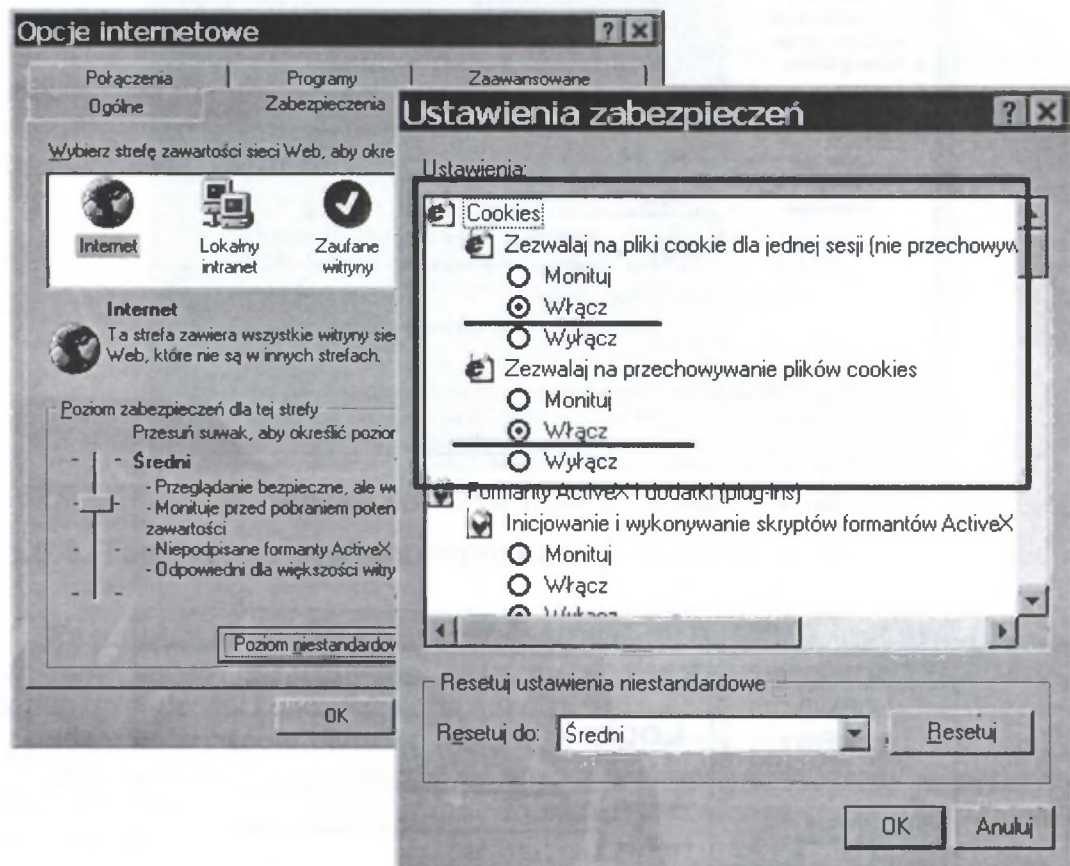


Rys. 42. Strona rejestracyjna systemu wprowadzania danych WWW-ISIS

Po udanym zarejestrowaniu się w systemie można zacząć korzystać z jego podstawowych funkcji (które wyświetlone są w lewej ramce, np. rys. 42):

- ADD** New record (dodawanie nowego rekordu)
- UPDATE** (uaktualnianie rekordów)
- SET DEFAULTS** (ustawianie wartości domyślnych)

W momencie wprowadzenia nowego rekordu należy sprawdzić opcje przeglądania stron (w szczególności należy sprawdzić czy nowo wprowadzone w polach wartości pozostają, kiedy przeglądamy poszczególne strony naszego formularza). W przypadku zaobserwowania nieprawidłowości zaleca się sprawdzenie parametrów dotyczących „cookies” i zezwolenie na ich tworzenie (patrz rys. 41 oraz 43).



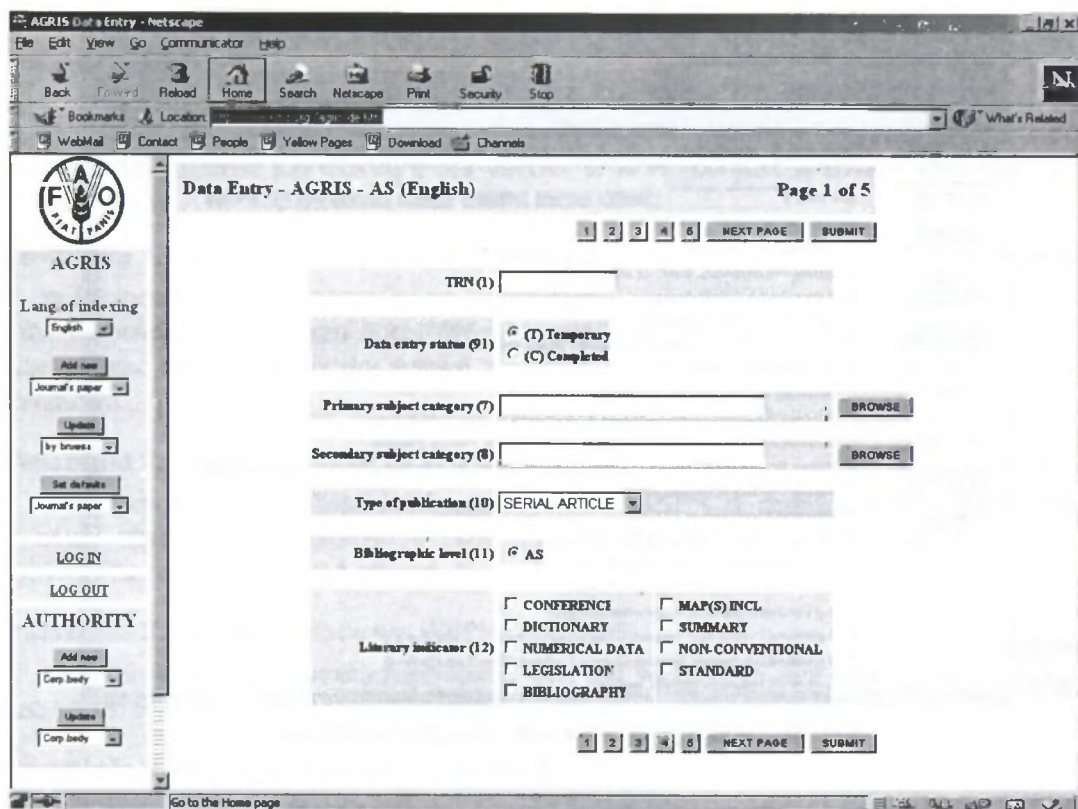
Rys. 43. Ustawienia „cookies” w Internet Explorer



REGUŁY WPROWADZANIA DANYCH

- Zamiast nawiasów trójkątnych można użyć nawiasów klamrowych { }⁴
- Dla pól powtarzalnych można użyć średnika ;
- Jeśli dla pola powtarzalnego widzimy tylko jedno wystąpienie pola tekstowego lub pola typu „text area”, to w celu odseparowania kolejnych wystąpień tego pola należy użyć średnika.

⁴ Aczkolwiek w wersji obecnej system WWW-ISIS nie robi użytku z takiej informacji.



Rys. 44. Przykładowe pole interfejsu do wprowadzania danych WWW-ISIS

Strony

Możliwe jest przemieszczanie się pomiędzy stronami używając klawiszy **NEXT PAGE**, **PREVIOUS PAGE**, oraz klawiszy z numerem strony (rys. 44).

Przycisk BROWSE (przeglądanie danych)

Jeśli przyciśniemy przycisk z napisem **BROWSE**, to możemy zobaczyć tablicę z podglądem danych (rys. 45), z której można wybrać wyświetlone wartości (zaznaczając pola typu „check”). Lista ta zaopatrzona jest w przyciski nawigacyjne (**BROWSE**, **NEXT PAGE**) oraz przyciski powrotu – **APPEND** do dodania zaznaczonego terminu do pola, **REPLACE** do zamiany zawartości pola z zaznaczonym terminem, **CANCEL** do powrotu do trybu edycji pola bez wprowadzania żadnych zmian.

W przypadku pól powtarzalnych obok terminów wyświetlone będą pola typu „check” (dające możliwość wyboru wielu terminów jednocześnie), zaś w przypadku pól niepowtarzalnych widoczne będą pola typu „radio” (dające możliwość pojedynczego wyboru).



Rys. 45. Tablica podglądowa dla pola „Corporate author”

Przyciski ADD GROUP

Jeśli pola ISIS zawierają podpola, zaś te mogą być powtarzalne, to podpola takie w interfejsie wprowadzania danych przedstawia się jako grupy powtarzalne **REPEAT** lub **ADD GROUP** (rys. 46). W momencie przycięnięcia przycisku **REPEAT** ukaże się następne wystąpienie pola wraz z podpolami do wprowadzenia danych.

Analytic material (AM, AMS)

Page 2 of 8

PREVIOUS PAGE | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | NEXT PAGE | SAVE RECORD |

Author's name & initials (28^a) Babacki J

Role (28^b) |

Author's name & initials (28^a) |

Role (28^b) |

Corporate Author American Arbitration Association

Paper Title (330) Another paper title for testing from the same | BROWSE |

Paper Title - Translation | BROWSE |

ADD GROUP

BROWSE | AD

Rys. 46. Grupy powtarzalne w interfejsie wprowadzania danych WWW-ISIS

3.7. Ćwiczenia



Przedstawione poniżej ćwiczenia wymagają zrozumienia środowiska sieci oraz zaznajomienia się obsługą serwerów WWW, co jest w miarę szczegółowo omówione w Aneksie E.

Ćwiczenie 1. Zagadnienia sieciowe (Aneks E)

TCP/IP

W komputerze nie podłączonym do sieci Internet zainstalować wirtualny system nazw domenowych w zbiorze hosts (www.isis.org, www.isis1.org, www.isis3.org). Przetestować fikcyjne nazwy domenowe z wiersza poleceń komendą ping „nazwa komputera”.

Ćwiczenie 2. Instalacja serwera WWW (Aneks E)

TCP/IP
WWW FTP

Zainstalować serwer XITAMI konfigurując katalog główny w serwera WWW w nowo utworzonym podkatalogu „C:\Temp\Webserver”, zaś serwera ftp „C:\Temp\ftprot”, przetestować serwery.

Ćwiczenie 3. Wirtualne serwery WWW (Aneks E)

TCP/IP
WWW

Na bazie wybranego serwera zdefiniować kilka tzw. wirtualnych serwerów, sprawdzić ich działanie.

Ćwiczenie 4. Obsługa serwerów (Aneks E)

TCP/IP
FTP

Jeśli pracujesz w lokalnej sieci komputerowej, postaraj się uruchomić usługę FTP, a następnie umieść w odpowiednim katalogu dowolne zbiory i poproś, aby inny użytkownik sieci ściągnął na swój komputer zbiory z twojego serwera.

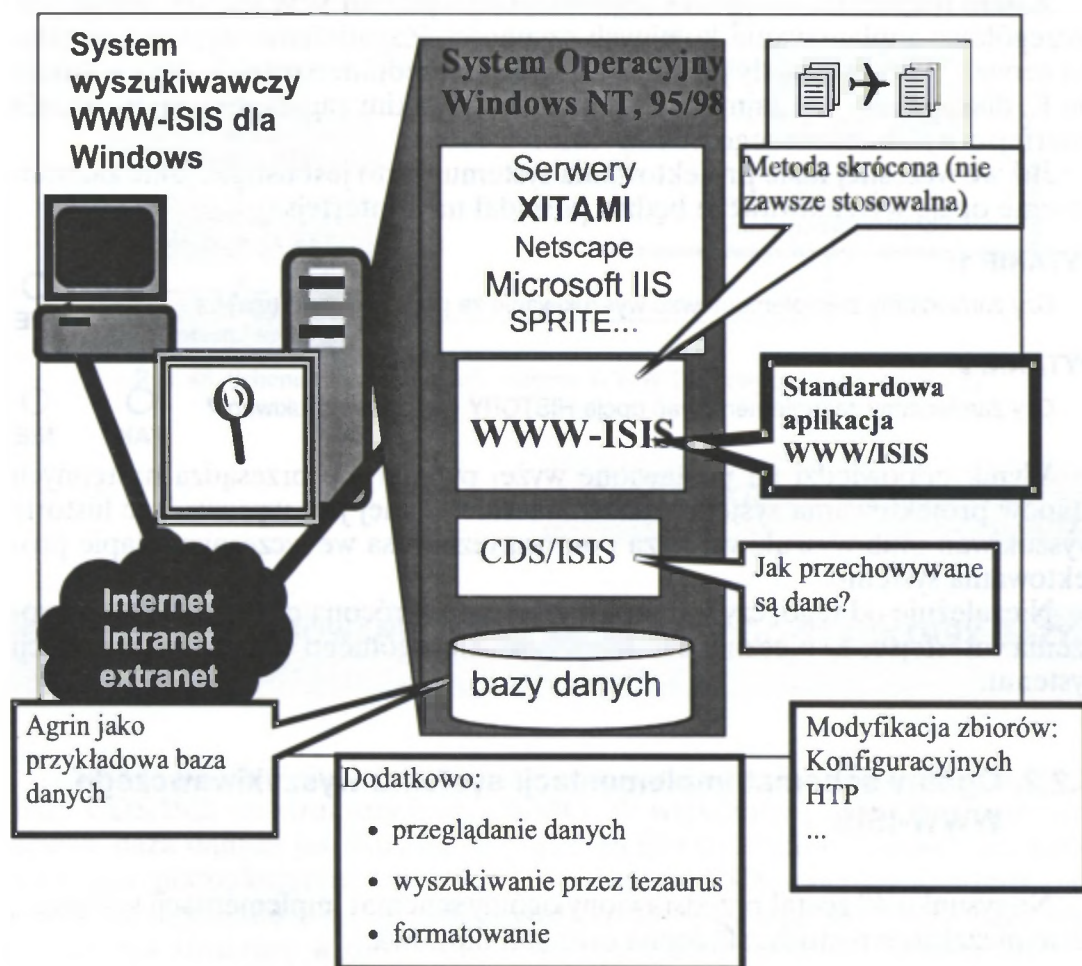
4. Wyszukiwanie w WWW-ISIS

4.1. Wstęp

Ta część poświęcona jest metodzie udostępniania bazy danych ISIS w sieci Internet/Intranet lub na CD-ROM wyłącznie w celu ich przeszukiwania. Po zapoznaniu się z tym rozdziałem użytkownik powinien być w stanie utworzyć w systemie WWW-ISIS swój własny interfejs wyszukiwawczy dla istniejących baz ISIS.

Zakładamy, że czytelnik zapoznał się z poprzednimi rozdziałami dotyczącymi samego systemu informacyjnego CDS/ISIS oraz, że posiada podstawową wiedzę na temat języka HTML.

Załączona wersja DEMO systemu jest stworzona w języku angielskim i przyjęliśmy ją jako przykładową bazę danych dla omawianego systemu wyszukiwawczego WWW-ISIS. Nie zakładamy jednak biegłej znajomości języka angielskiego u czytelnika, dlatego, gdzie tylko to możliwe, będzie występowało polskie tłumaczenie.



Rys. 47. Omówienie interfejsu wyszukiwawczego WWW-ISIS

Podsystem wyszukiwawczy WWW-ISIS jest częścią systemu WWW-ISIS i dzięki możliwości modyfikacji odpowiednich zbiorów tekstowych jest łatwy w implementacji nawet dla początkujących użytkowników, zaś jego bogate możliwości oraz stosunkowo krótki czas konieczny do stworzenia w pełni profesjonalnego narzędzia, powinny zadowolić nawet najbardziej wymagających użytkowników.

W rozdziale przedstawione są dwie metody implementacji systemu szybka – choć nie zawsze stosowalna oraz metoda standardowa. Do pełnej implementacji systemu może okazać się konieczne zmodyfikowanie zbioru konfiguracyjnego, zbiorów html oraz plików http. Omawiamy to w dalszej części rozdziału. Całość zamykają dodatkowe informacje dotyczące m.in. zaimplementowania tezaursu, wyszukiwania w wielu bazach jednocześnie oraz eksportu danych.

4.2. Plan działania

4.2.1. Podstawowe pytania

Zanim rozpoczniemy proces implementacji systemu WWW-ISIS, ważne jest szczegółowe zaplanowanie kolejnych czynności. Zagadnienia sieciowe, instalacja serwera WWW zostały już omówione w poprzednim rozdziale oraz w Aneksie E, dlatego obecnie zajmiemy się przede wszystkim zagadnieniami tworzenia interfejsu wyszukiwawczego WWW-ISIS.

Już we wczesnej fazie projektowania systemu warto jest ustalić, jakie zaawansowane opcje wyszukiwawcze będzie posiadał nasz interfejs.



PYTANIE 1:

Czy zamierzamy zaimplementować wyszukiwanie za pomocą tezaursu?¹

☐ TAK

☐ NIE



PYTANIE 2:

Czy zamierzamy zaimplementować opcję HISTORY (historia wyszukiwań)?

☐ TAK

☐ NIE

Wynik odpowiedzi na postawione wyżej pytania nie przesądza następnych etapów projektowania systemu, choć znacznie łatwiej jest wprowadzać historie wyszukiwań i/lub wyszukiwanie za pomocą tezaursu we wczesnym etapie projektowania systemu.

Niezależnie od tego, czy wybierzemy metodę skróconą czy standardową tworzenia interfejsu, konieczne jest naszkicowanie ogólnego planu implementacji systemu.

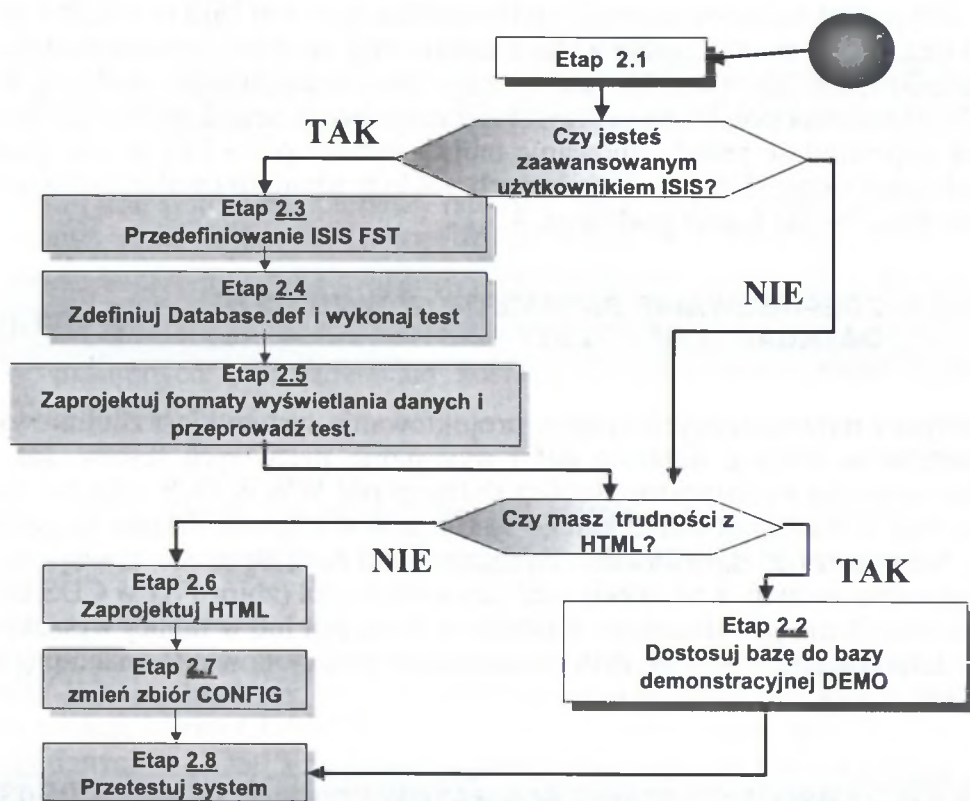
4.2.2. Ogólny schemat implementacji systemu wyszukiwawczego WWW-ISIS

Na rysunku 48 został przedstawiony ogólny schemat implementacji systemu – dwie niezależne metody: skrócona oraz standardowa.

¹ Może okazać się zbyt trudne w pierwszym podejściu.

ETAP 2.1. INSTALACJA SYSTEMU

Pierwszym proponowanym krokiem jest zainstalowanie samego systemu WWW-ISIS. Poszczególne etapy zostały omówione w rozdziale 3 oraz w Aneksie E.



Rys. 48. Schemat implementacji systemu WWW-ISIS (wyszukiwanie danych)

METODA SKRÓCONA



ETAP 2.2. DOSTOSOWANIE ISTNIEJĄCEJ BAZY DO STRUKTURY BAZY DEMO

„Szybka ścieżka” implementacji systemu polega na dostosowaniu istniejącej bazy CDS/ISIS do struktury bazy DEMO. W większości przypadków, gdy wyjściowa baza danych jest bazą bibliograficzną jest to możliwe. Sposób ten polecany jest początkującym użytkownikom systemu ISIS, oczywiście wtedy, gdy mamy do czynienia z bazami bibliograficznymi – gwarantuje to, że baza wyjściowa ma strukturę w dużej mierze kompatybilną ze strukturą bazy DEMO (z dokładnością do przenumeroowania znaczników pól). Minimum wiedzy, jakie jest wymagane, ogranicza się w takim przypadku do „przetłumaczenia” oryginalnego

nalnej bazy tak, aby jej struktura była zgodna ze strukturą bazy DEMO. Szczegółowa procedura omówiona jest w podrozdziale 4.4.



ETAP 2.3. PRZEDDEFINIOWANIE ZBIORU FST

**CDS/ISIS
: FST**

W przypadku zaawansowanych użytkowników systemu ISIS oraz/lub w przypadku baz, które nie są zgodne z standardem bazy DEMO, konieczne staje się przeddefiniowanie zbioru FST, co wymaga przeanalizowania struktury bazy danych, określenia pól, które w przyszłości mają być przeszukiwane. Konieczne jest też odpowiednie przeddefiniowanie tablicy wyboru pól – FST w taki sposób, aby indeksy z wszystkich przeszukiwanych pól były identyfikowalne przez odpowiednie przedrostki (patrz podrozdz. 4.5.1).



ETAP 2.4. ZDEFINIOWANIE PARAMETRÓW W ZBIORZE DATABASE.DEF+ TESTY

WWW-ISIS

Jednym z najważniejszych etapów projektowania systemu jest zdefiniowanie parametrów w zbiorze database.def i wykonanie pierwszych testów. Jak już zostało wcześniej wspomniane, tablica definicji pól WWW-ISIS odgrywa decydującą rolę w funkcjonowaniu całego systemu. Jest odpowiedzialna za generowanie formularzy z zapytaniem. Tablice wyszukiwań są ściśle powiązane ze zbiorem odwróconym, a właściwie z tablicą wyboru pól (zbiór FST w CDS/ISIS). Błędne określenia przedrostków w tablicy wyboru pól lub w tablicy wyszukiwań (zbiór database.def w WWW-ISIS) uniemożliwi funkcjonowanie systemu (patrz podrozdz. 4.5.2).



ETAP 2.5. ZAPROJEKTOWANIE FORMATÓW WYŚWIETLANIA DANYCH + TESTY

**CDS/ISIS
FORMATY**

Etap ten polega na zaprojektowaniu formatów wyświetlania danych i przeprowadzeniu testów. Jest to dość czasochłonny proces, a jego rezultaty zależą od biegłości w posługiwaniu się językiem formatowania CDS/ISIS. Ponadto sposób, w jaki będą one wyświetlane, zależy w dużej mierze od znajomości języka HTML. W podrozdziale 4.5.3 znajduje się kilka przykładów oraz wskazówek, jak projektować właściwe formaty.



ETAP 2.6. ZAPROJEKTOWANIE HTML+ TESTY

<HTML>

Kolejny etap projektowania systemu polega na zaprojektowaniu HTML. W przypadku słabej znajomości HTML zaleca się wprowadzanie nieznacznych zmian w istniejących zbiorach HTML. Najważniejsze jest jednak odpowiednie ustawienie połączeń hipertekstowych między nowymi formatami a odwołaniami do baz danych.



Ustawienie i sprawdzenie ustawień w zbiorze isis3w.exe.cfg. Czynność ta wymaga co najwyżej niewielkich zmian w zbiorze konfiguracyjnym (patrz Aneks A).

ETAP 2.8. TESTOWANIE SYSTEMU



Nawet jeśli system na „pierwszy rzut oka” wygląda poprawnie, zaleca się jego testowanie przez kilka dni i sprawdzanie wszelkich formularzy wyświetlania, połączeń, szczególnie w przypadkach interfejsów wielojęzycznych. Rzecz jasna, etap ten wymagany jest niezależnie od trybu zaimplementowania systemu (szybki i standardowy).

W następnych podrozdziałach zostaną bardziej szczegółowo omówione poszczególne kroki.

4.3. Podstawowe kroki tworzenia systemu wyszukiwawczego WWW-ISIS

W podrozdziale tym przedstawimy proces tworzenia własnej aplikacji WWW-ISIS i zakładamy, że czytelnik opanował materiał dotyczący podstawowych funkcji Micro CDS/ISIS. W części 2 omówiono m.in. sposób przechowywania danych w CDS/ISIS. Jego przypomnienie może okazać się niezbędne do zrozumienia poszczególnych etapów tworzenia systemu.

Pragnąc dokonać implementacji interfejsu WWW-ISIS w istniejącej bazie ISIS projektant powinien pamiętać, że możliwości przeszukiwania bazy zależą od ustalenia tzw. punktów dostępu, czyli ustalenia zbioru indeksującego dla bazy ISIS, czyli tzw. zbioru odwróconego, co dokonuje się dzięki ustaleniu odpowiednich przedrostków w FST.

KROK 1 – USTAL, KTÓRE ELEMENTY BĘDĄ WYSZUKIWANE



Już w momencie projektowania bazy ISIS winniśmy ustalić, które elementy będą w przyszłości przeszukiwane. Na przykład dla typu rekordu ARTYKUŁ (PAPER) ustalamy następujące elementy:

- *Autor artykułu* (Author),
- *Tytuł* (Title),
- *Tytuł czasopisma* (Journal Title),
- *Data publikacji* (Publication Date),

- *Deskryptory* (Descriptors),
- *Słowa kluczowe z abstraktu* (Keywords From Abstract) itd.

Naturalnie dla zbyt długich pól, jak na przykład „Tytuł”, należałoby ograniczyć elementy wyszukiwawcze tylko do słów kluczowych.



KROK 2 – USTAL ODPOWIEDNIE PREFIKSY W FST

CDS/ISIS
FST

Mając na względzie zasadę odróżniania pól wyłącznie przez prefiksy (przedrostki), należy ustalić odpowiednie przedrostki dla przeszukiwanych pól. Prefiks taki winien poprzedzać wszystkie te dane, które chcemy wydobyć z danego pola, co oznacza, że projektowaniu tablicy wyboru pól (FST) należy poświęcić szczególną uwagę.



KROK 3 – ZDEFINIUJ TABLICĘ DEFINICJI WYSZUKIWAŃ (DATABASE.DEF...,)

WWW-ISIS

Na tym etapie winniśmy określić tablicę definicji przeszukiwań w WWW-ISIS, czyli zbiory database.def.... oraz browse.def... Przechowują one informację na temat wszystkich przeszukiwanych pól i przyporządkowanych im prefiksów dla dowolnych języków etnicznych (dla angielskiego – database.def.eng, browse.def.eng; czy dla polskiego – database.def.pol, browse.def.pol itd.).



KROK 4 – ZDEFINIUJ FORMATY WYŚWIETLANIA DANYCH W HTML + ZBIORY HTP (OPCJONALNIE)

CDS/ISIS <HTML>
FORMATY HTP

Należy zdefiniować formaty wyświetlania danych w HTML. Zwykle należy przygotować uproszczony format do wyświetlania wybranych danych, a następnie pełny format do wyświetlania wszystkich danych z rekordu zgodnie z tablicami definicji wyszukiwań. Ponadto może okazać się konieczne przeddefiniowanie odpowiednich szablonów htp dla poszczególnych języków etnicznych (co może mieć miejsce w przypadku tworzenia interfejsu wielojęzycznego).



KROK 5 – PRZEGLĄDANIE DANYCH (BROWSE.DEF...)

WWW-ISIS

Funkcja przeglądania danych umożliwia początkującym użytkownikom systemu przeglądanie danych w określonym porządku. Na przykład w przypadku danych bibliograficznych, możemy przyjąć, że dane dotyczące artykułów mogą być przeglądane w następującym porządku:

- Autor
 - Tytuł artykułu

Zaś układ przeglądania książek będzie następujący:

- Autor
 - Tytuł
 - Numer ISBN

Za przeglądanie danych odpowiedzialne są zbiory browse.def... Przechowują one informację na temat wszystkich przeglądanych pól i przyporządkowanych im prefiksów dla dowolnych języków etnicznych (dla angielskiego – browse.def.eng; czy dla polskiego – browse.def.pol itd.)

PAMIĘTAJ

O właściwym przeprojektowaniu tablicy wyboru pól (FST) w ISIS (ustalenie właściwych przedrostków) ze względu na wyszukiwanie oraz przeglądanie danych w systemie WWW-ISIS.



4.4. Metoda skrócona implementacji systemu WWW-ISIS

Ten sposób zaimplementowania systemu polega w zasadzie na „przetłumaczeniu” istniejącej bazy CDS/ISIS tak, aby jej struktura była zgodna z bazą DEMO systemu WWW-ISIS. Naturalnie proces ten nie zawsze jest możliwy, gdyż z oczywistych powodów nie każda baza danych CDS/ISIS jest bazą bibliograficzną. Zanim zdecydujemy się na zastosowanie metody skróconej należy odpowiedzieć sobie na poniższe pytania.

METODA SKRÓCONA



PYTANIE 1:

Czy mamy do czynienia z bibliograficzną bazą danych opisującą artykuły, książki itd.?

☐ TAK

☐ NIE

PYTANIE 2:

Czy jesteśmy w stanie dokonać takiej modyfikacji tablicy wyboru pól (FST), aby była zgodna z tablicą FST bazy DEMO?

☐ TAK

☐ NIE

PYTANIE 3:

Czy potrafimy dokonać eksportu oraz importu danych wykorzystując system ISIS (wersje DOS lub Windows)?

☐ TAK

☐ NIE

W przypadku odpowiedzi pozytywnych na powyższe pytania możemy przystąpić do implementacji systemu metodą skróconą.

Do podstawowych korzyści płynących z zastosowania metody skróconej należą:

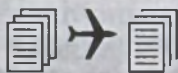
- Brak konieczności definiowania formatów wyświetlania danych oraz zaawansowanej wiedzy języka HTML (gdyż struktura naszej bazy jest zgodna ze standardami bazy DEMO).
- Brak konieczności przeddefiniowania zbioru FST.
- Brak konieczności przeddefiniowania tablicy wyszukiwań.

Istnieją jednak wady skróconej metody implementacji WWW-ISIS. Przystosowana baza danych nie jest taką samą bazą, jaką była przed modyfikacją². i nawet, jeśli system będzie pełnił dość ograniczoną rolę³, tłumaczenie bazy tylko dla celów utworzenia interfejsu wyszukiwawczego WWW-ISIS może być, co najmniej dziwne. Bowiem za każdym razem, kiedy będziemy chcieli dokonać uaktualnienia bazy przetłumaczonej do standardu bazy DEMO, konieczne stanie się ponowne przetłumaczenie części bazy, a następnie dokonanie aktualizacji samego zbioru odwróconego oraz całego systemu wyszukiwawczego.

Mając na względzie argumenty „za” oraz „przeciw” zastosowaniu metody skróconej lub metody standardowej należy poważnie zastanowić się nad wyborem tak, aby w przyszłości nakład pracy związany z utrzymaniem bazy nie przeszył korzyści z zastosowania systemu WWW-ISIS.



METODA SKRÓCONA



UWAGA:

Nie zaleca się wyboru metody skróconej w przypadku baz innych niż bibliograficzne!

Przedstawiona struktura bazy DEMO jest identyczna ze strukturą istniejącej bazy katalogu bibliotecznego należącej do FAO⁴. W tabeli porównawczej dla bazy DEMO (tab. 3) została przedstawiona lista wszystkich pól bazy DEMO. W kolumnie „Wyjaśnienie” znajduje się wyjaśnienie podanych pól.

² Nawet, jeśli dla potrzeb interfejsu wyszukiwawczego WWW wykorzystuje się tę samą strukturę, to dla interfejsu wprowadzania danych, zaleca się utworzenie repliki bazy w sieci.

³ Np. w intranecie – czyli w obrębie sieci lokalnej wykorzystującej technologię internetową.

⁴ Food and Agriculture Organization.

Tabela 4

Tabela porównawcza dla bazy DEMO

Opis pola (ang./ (tłum. pol.))	Podpole	Pole	Powtarzalność	Wyjaśnienie	Pole w tłumaczonej bazie	Powtarzalność
1	2	3	4	5	6	7
Abstract . [abstrakt]		79				
AGRIIS/CARIS categories (rough classification) [klasyfikacja, opis skrócony]		8	T	Klasyfikacja może być zastosowana do podania kodów klasyfikacji		
AGRIIS/CARIS categories full (rough classification) [klasyfikacja, opis pełny]		810		Może oznaczać pełne wytłumaczenie klasyfikacji rekordów		
Bibliographic level (rodzaj opisu bibliograficznego)		11	N	M oraz MS dla monografii AM oraz AMS dla artykułów z książek AS dla artykułów z czasopism		
Collation - Monographic level		56	N			
Collation - Serial level		57	N			
Conference date [data konferencji]		50	N			
Conference name(s) and number [nazwa i numer konferencji]	^e^f^s^o^n	46	N	Podpola ^e^f^s^o są przeznaczone odpowiednio dla języków: angielskiego, francuskiego, hiszpańskiego zaś podpole ^n pełni rolę oznaczenia numeru konferencji		
Conference place [miejsce konferencji]		48	N			
Corporate Author - Monographic level [autor zbiorowy monografii]	^e^f^s^o	31	T	Podpola są dla odpowiednich wersji językowych: angielskiej, francuskiej, hiszpańskiej itd.		
Corporate Author Analytic level [autor zbiorowy artykułu z monografii]	^e^f^s^o	30	T	Podpola są dla odpowiednich wersji językowych: angielskiej, francuskiej, hiszpańskiej itd.		
Date of publication - M lev. [data publikacji – poziom M]		54	N	Data publikacji np. Kwiecień 1997		
Date of publication - S lev. . [data publikacji – poziom S]		55	N	jw.		
Division symbol [symbol działu]		19	T	Zbyt szczegółowe		
Document class code [kod klasyfikacji dokumentu]		13	N			
Document JOB NO.		20	N	Zbyt szczegółowe		
Edition number		51	N			
FAOBIB Accession number [identyfikator rekordu]		1	N	W tłumaczonej bazie powinno się określić identyfikator rekordu		
File assignment		21	N	Zbyt szczegółowe		
Form of document . [forma dokumentu]		72	T			
General source		67	N			
Geographic subject		75	T	Każdy deskryptor jest umieszczony w oddzielnych wystąpieniach pola		

1	2	3	4	5	6	7
Holding library [<i>informacja, czy dokument jest w posiadaniu biblioteki</i>]		14	T	Dość powszechne w przypadku kilku bibliotek posiadających wspólny katalog		
Institution acronyms . [<i>akronim instytucji</i>]		70	N			
ISBN		26	N			
ISSN		27	N			
Language full name		430	T			
Language full name [<i>pełna nazwa języka etnicznego</i>]		440	T			
Language of text - A lev. [<i>język tekstu – poziom A</i>]		43	T			
Language of text - M lev.		44	T			
Library call number [<i>przyp. sygnatura biblioteki</i>]		22	N			
Library class. codes		87		Zbyt szczegółowe		
Literary indicator [<i>informacja biblioteczna</i>]		12	T	Opis typu publikacji Możliwe wartości: CONFERENCE (<i>materiały konferencyjne</i>) DICTIONARY (<i>słowniki</i>) NUMERICAL DATA (<i>dane numeryczne</i>) LEGISLATION (<i>zagadnienia prawne</i>) BIBLIOGRAPHY (<i>bibliografia</i>) MAP(S) INCL. (<i>zawierają mapy</i>) SUMMARY (<i>streszczenia</i>) NON-CONVENTIONAL `(<i>nie konwencjonalne</i>) STANDARD (<i>standard</i>)		
Local descriptors		78				
Microfiche number . [<i>numer mikrofilmu</i>]		65	N			
Notes - Analytic level		58	N			
Notes - Monographic level		59	N			
Personal Author - Analytic level [<i>autor artykułu</i>]	^a^b	28	T			
Personal Author - Monographic level [<i>autor monografii</i>]	^a^b	29	T			
Place of publication [<i>miejsce konferencji</i>]		52	N			
Publication year . [<i>rok publikacji</i>]		64	N	Tylko 4 cyfry np. 1997		
Publisher [<i>wydawca</i>]		53	N			
Related lang. Versions	^e^f^s^o	60	N	Połączenia do numeru dostępu. Podpola są dla odpowiednich wersji językowych: angielskiej, francuskiej, hiszpańskiej itd.		

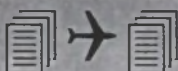
1	2	3	4	5	6	7
Report number - Analytic level [<i>numer raportu poziom analityczny</i>]		17	N			
Report number - M level [<i>numer raportu poziom monografii</i>]		18	N			
Secondary title element(s) - Analytic level	^e^f^s^o^t	35	N	Podpola są dla odpowiednich wersji językowych: angielskiej, francuskiej, hiszpańskiej itd.		
Secondary title element(s) Monographic level	^e^f^s^o^t	36	N	Podpola są dla odpowiednich wersji językowych: angielskiej, francuskiej, hiszpańskiej itd.		
Serial title(s) [<i>tytuł czasopism(a)</i>]	^e^f^s^o	41	N	Podpola są dla odpowiednich wersji językowych: angielskiej, francuskiej, hiszpańskiej itd.		
Thesaurus secondary index terms		73	N	Deskryptory umieszcza się w nawiasach kwadratowych		
Thesaurus subject terms		74	T	Każdy deskryptor jest umieszczony w oddzielnych wystąpieniach pola		
Title(s) - Analytic level [<i>autor zbiorowy czasopisma</i>]	^e^f^s^o^t	33	N	Podpola są dla odpowiednich wersji językowych: angielskiej, francuskiej, hiszpańskiej itd.		
Title(s) Monographic level [<i>tytuł monografii</i>]	^e^f^s^o^t	34	N	Podpola są dla odpowiednich wersji językowych: angielskiej, francuskiej, hiszpańskiej itd.		

Przyjrzyjmy się bliżej omawianej metodzie. Zakładamy, że czytelnik dysponuje podstawową wiedzą na temat opisu bibliograficznego oraz rządzących nim zasad. Dla przypomnienia podajemy uproszczony sposób opisu bibliograficznego („description level”) – analytic (*artykuł w czasopiśmie*), monographic (*publikacja zwarta*), serial (*czasopismo*).

Jeśli rekord opisuje **książkę**, to opis bibliograficzny jest określony jako M lub MS. W takim rekordzie pola AUTHOR, CORPORATE AUTHOR, TITLE, oraz TITLE SUPPLEMENT są określone jako „monographic level” („poziom monograficzny”). Ponadto, jeśli opis został zakwalifikowany jako MS, to w danym rekordzie powinny pojawić się również dane dotyczące opisu czasopism „Serial Level” (tytuł, numer itd.).

Jeśli rekord opisuje **artykuł z książki** (wydawnictwa zwarte), to rodzaj opisu bibliograficznego powinien zostać określony jako AM lub AMS. W tego typu rekordach powinny być wypełnione następujące pola: AUTHOR, CORPORATE AUTHOR, TITLE, oraz TITLE SUPPLEMENT (czyli pola na znacznikach 28, 30, 33, 35). W rekordach tych pojawiają się też pola dotyczące rekordów typu Monograph (*publikacje zwarte*) w celu ustalenia książki, z której pochodzi dany artykuł (tytuł książki, autorzy itd.).

Jeśli rekord opisuje **artykuł z czasopisma**, to rodzaj opisu bibliograficznego powinien zostać określony jako AS. W tego typu rekordach powinny być wypełnione następujące pola: AUTHOR, CORPORATE AUTHOR, TITLE, oraz TITLE SUPPLEMENT (pola oznaczone jako 28, 30, 33, 35). W rekordzie mogą pojawić się informacje dotyczące rekordów typu SERIAL w celu ustalenia, z jakiego czasopisma jest dany artykuł (tytuł czasopisma i jego numer itd.). Po zapoznaniu się ze strukturą bazy DEMO, możemy przeprowadzić procedurę translacyjną.

**KROK 1 – PORÓWNANIE BAZY TŁUMACZONEJ Z BAZĄ DEMO**

W ostatniej kolumnie w tabeli porównawczej (tab. 4) należy wprowadzić odpowiedni znacznik z tłumaczonej bazy korespondujący z danym polem w bazie DEMO. Zaleca się rozpoczęcie procesu tłumaczenia od znaczników pól, co do których możemy bez problemu ustalić zawartość pól: Autor (DEMO=Author), Autor zbiorowy (DEMO=Corporate Author), Tytuł (DEMO=Title).

Należy podkreślić, że istnieje niewielka liczba pól, które należy wziąć pod uwagę w całym procesie tłumaczenia:

- *Analytic Title* (Tytuł artykułu),
- *Monographic Title* (Tytuł monografii),
- *Analytic individual authors* (Autorzy artykułów),
- *Monographic individual authors* (Autorzy wydawnictw zwartych),
- *Analytic Corporate Authors* (Autorzy zbiorowi artykułów),
- *Conference data* (Data konferencji),
- *Descriptors* (Deskryptory),
- *Document type* (Typ dokumentu),
- *Year of publication* (Data publikacji).

Jednak w przypadku tłumaczenia pozostałych pól zaleca się daleko idącą ostrożność.

**UWAGA:**

Mogą pojawić się pewne problemy w przypadku pól mających podpole: ^e, ^f, ^s, ^o i w niektórych przypadkach pól z podpolem ^t, które oznaczają odpowiednio języki: angielski, francuski, hiszpański itd. Podpole te odnoszą się do tytułów, nazw złożonych, tytułów czasopism, tytułów konferencji. Nawet, jeśli w naszej bazie nie zamierzamy odróżniać różnych wersji językowych, to powinno się zachować te podpole. Np. jeśli w naszej bazie będzie znajdowało się pole opisujące oryginalny tytuł oraz pole, w którym będzie jego tłumaczenie na język angielski, to należy „przetłumaczyć” pole *tytuł oryginalny* na ^o, zaś to podpole, w którym jest przechowywane tłumaczenie, winno mieć oznaczenia ^t.

Przetłumaczenie pozostałych pól nie powinno nastroczać trudności.

**KROK 2 – TŁUMACZENIE TABLICY SELEKCJI PÓL (FST)****CDS/ISIS
FST**

Po ustaleniu przyporządkowania pól z tłumaczonej bazy, bazie DEMO, należy przystosować tablicę wyboru pól (zbiór FST). W procesie tłumaczenia należy pamiętać o zachowaniu powtarzalności pól (zgodnie z bazą DEMO).

KROK 3 – TWORZENIE ZBIORU ISO2709**CDS/ISIS**

W kolejnym kroku należy uruchomić program CDS/ISIS (wersja DOS lub Windows) i dokonać eksportu naszej bazy do standardu ISO2709, przy zachowaniu przeprojektowanego (przetłumaczonego w kroku 2) zbioru FST (tablicy selekcji pól). Należy upewnić się, czy ogranicznik pól oraz rekordów nie pokrywają się z zawartością któregośkolwiek z rekordów. Zaleca się wykorzystanie znaków ~ oraz & jako odpowiednio ograniczników pól oraz rekordów.

KROK 4 – ZAŁADOWANIE DANYCH DO STANDARDU DEMO**CDS/ISIS**

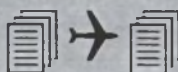
W kolejnym i przedostatnim kroku należy zainicjować pustą bazę DEMO w środowisku ISIS, korzystając z tego samego lub osobnego folderu, np. DEMO. Zakładamy, że czytelnik rozumie sposób konfigurowania zbioru syspar.par w CDS/ISIS (por. rozdz. 2 i Aneks B) tak, aby jego baza była „widoczna” w środowisku ISIS. Następnie należy załadować zbiór otrzymany w kroku 2. Jeśli operacja nasza nie zakończyła się powodzeniem, oznacza to prawdopodobnie, że przy eksporcie danych nastąpiły problemy z właściwym oznaczeniem separatorów pól i rekordów. W takim przypadku zaleca się ponowne przeprowadzenie procesu eksportu naszej bazy i wykorzystanie innych separatorów np., | oraz #. Zanim jednak tego dokonamy, powinniśmy sprawdzić występowanie znaków separatorów w wyeksportowanym poprzednio zbiorze MST.ISO (za pomocą odpowiedniego edytora tekstowego).

KROK 5 – PRZETESTOWANIE BAZY ZGODNEJ ZE STANDARDEM DEMO

W ostatnim kroku należy dokonać sprawdzenia działania nowej bazy zgodnej ze standardem DEMO. Do istotnych pytań, na jakie należy odpowiedzieć należą:



METODA SKRÓCONA



PYTANIE 1:

Czy rozmiar bazy jest taki sam jak oryginalnej?⁵

☐
TAK

☐
NIE

PYTANIE 2:

Czy nowe pola są właściwie przetłumaczone?

☐
TAK

☐
NIE

PYTANIE 3:

Czy zostały zachowane podpola?

☐
TAK

☐
NIE

PYTANIE 4:

Czy zachowano powtarzalność podpól (istotne w przypadku pól, które mają być przeszukiwane)?

☐
TAK

☐
NIE

PYTANIE 5:

Czy deskryptory w polach mają taką samą strukturę?

☐
TAK

☐
NIE

W celu przejrzania wszystkich rekordów należy użyć pełnego formatu wyświetlania danych (opcja FORMAT). Jeśli nasz test wypadnie niepomyślnie, należy dokonać ponownego tłumaczenia tablicy wyboru pól i powtórzyć kroki 3-5.

Jeśli zawartość bazy nie budzi zastrzeżeń, w kolejnym etapie należy:

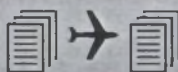
- Zmienić nazwę oryginalnych zbiorów bazy DEMO w katalogu, gdzie znajduje się oprogramowanie WWW-ISIS, poza zbiorami samej bazy danych (podczas instalacji ustaliliśmy lokalizację odpowiednich katalogów). Zbiory, których nazwy należy zmienić to: demo.mst, demo.xrf, demo.ifp, demo.n01, demo.n02, demo.l01, demo.l02.
- Skopiować zbiory bazy DEMO do katalogu systemu WWW-ISIS, gdzie powinny się znajdować zgodnie z konfiguracją.
- Sprawdzić funkcjonowanie systemu przeglądarką Internet Explorer i/lub Netscape.

Poniżej zostanie pokazany proces szybkiego przystosowania do standardów bazy DEMO, przykładowej bazy danych AGRIN (jej pełny opis znajduje się w Aneksie F).

4.4.1. Przykład – Przystosowywanie bazy AGRIN do standardów bazy DEMO



METODA SKRÓCONA



Po stworzeniu tabeli (na wzór tab. 4), należy przygotować odpowiednią tablicę wyboru pól. Poniżej przedstawiona (tab. 5) jest gotowa tablica FST wraz z komentarzami (oryginalna tablica FST powinna być oczywiście pozbawiona komentarzy).

⁵ Jeśli w procesie tłumaczenia bazy wybraliśmy tylko jej podzbiór, to punkt ten można pominąć.

Tablica wyboru pól (FST) dla bazy AGRIN

Tabela 5

Wiersz FST	Komentarz
1 0 MFN (6)	
8 0 (v7/)	Chociaż nie wykorzystane dla celów wyszukiwawczych może okazać się przydatne w przyszłości.
11 0 V11	
12 0 (v12/)	
17 0 (v1300/)	
18 0 (v2300/)	
26 0 (v2320/)	
27 0 (4320/)	
28 6 '/^a/' if p(v1100) then V1100 fi	Każda wartość położona między nawiasami będzie poprzedzona przedrostkiem ^a i przechowywana jako osobne wystąpienie pola
29 6 '/^a/' if p(v2100) then V2100 fi	Jak wyżej
30 0 '/^e/' if p(v1110) then V1110/ fi	Prawdopodobnie nazwy autorów zbiorowych są po angielsku
31 0 '/^e/' if p(v2110) then V2110/ fi	Jak wyżej
33 0 "^e"v1200, ^f"v1202,"^s"v1204,"^o"v1206	Tytuł artykułu składa się z różnych wersji językowych
34 0 "^e"v2200 ^f"v2202,"^s"v2204,"^o"v2206	Tytuł książki składa się z różnych wersji językowych
41 0 "^e"V4230". "V4231	
430 0 (v1600/)	Tylko pole 430 jest wykorzystywane do wyszukiwań (w bazie DEMO jest to wartość zakodowana, wyszukiwanie będzie możliwe przez podanie kodu językowego, jeśli taki istnieje)
440 0 (v2600/)	Patrz wyżej
46 0 if p(v1210) then "^e"v1210 else ^e"v2210 fi	
48 0 if p(v1211) then "^e"v1211 else ^e"v2211 fi	
50 0 if p(v1213) then "^e"v1213 else ^e"v2213 fi	
51 0 (V2250/)	
52 0 (V2401/)	
53 0 (V2402/)	
54 0 (V2403/)	
55 0 (V4402/)	
56 0 V1400", ", V2500	
57 0 V4500	
58 0 V1610	
59 0 V2610	
73 0 V5502	Jeśli tytuły miałyby być po francusku lub hiszpańsku użylibyśmy odpowiednio pól o znacznikach V5522 lub V5542
74 2 V5501	Używamy techniki selekcji 2, co powoduje, że każdy deskryptor umieszczony w nawiasach staje się oddzielnym polem.
78 0 v5820	
79 0 V6860". ", V6861	

Po odpowiednim przeprojektowaniu tablicy FST (wyboru pól) należy dokonać odpowiedniego przetłumaczenia bazy do standardu DEMO. Rezultatem będzie zbiór zapisany w standardzie ISO2709, który winien być załadowany do

bazy ISIS, sprawdzony a następnie skopiowany do środowiska WWW-ISIS. Ponieważ pola V64 (year of publication) i V75 (Geographic Subject) nie zostały przetłumaczone należy odwołania do nich usunąć w zbiorze database.def.eng (w katalogu \demo):

```
# FAODOC database def file# created 8/2/1999 by MarS
[mons; MON; Selected publications]
[sort; @SCHR; Chronologically]
[sort; @SAUT; by Author]
[sort; @STIT; by Title]
[format; @ESMON; Short format]
[format; @EBMON; Bibliography format]
[format; @ELMON; Long format]
[format; @EFMON; Full format]
"Keywords from record", "keywords_record", "K=", "1" ;
"restricted_terms", "keywords_record"
[mona; MON; Selected publications]
[sort; @SCHR; Chronologically]
[sort; @SAUT; by Author]
[sort; @STIT; by Title]
[format; @ESMON; Short format]
[format; @EBMON; Bibliography]
[format; @ELMON; Long format]
[format; @EFMON; Full format]
"Keywords from Title", "mtitle", "TI=", "1" ;
"Personal Author", "mauthor", "A=", "1" ;
"Corporate Author", "mcorp_author", "C=", "1" ;
"Keywords from Conference", "mconf", "CO=", "1" ;
"Descriptors", "mthesaurus", "D=", "1" ;
"Geographic Coverage", "mgeo", "GE=", "1" ;
"Language", "mlanguage", "L=", "1" ;
"Document type", "mdoc_type", "DT=", "1" ;
"Year, e.g. 1981;1982", "myear", "YY=", "0" ;
"restricted_terms", "mtitle", "mauthor", "mcorp_author", "mconf",
"myear", "mgeo", "mthesaurus"
```

Następnie powinniśmy przetestować interfejs WWW w przetłumaczonej bazie. Za każdym razem, ilekroć pragniemy dokonać uaktualnienia, cały proces tłumaczenia bazy danych powinien być powtórzony.

4.5. Standardowy sposób implementacji systemu

W tej części prezentujemy szczegółowy opis poszczególnych faz standardowego sposobu implementacji interfejsu wyszukiwawczego WWW-ISIS. Sposób ten można zastosować dla dowolnej bazy ISIS. Podejście to obejmuje etapy 2.3 – 2.6 (por. podrozdz. 4.2.2).

Uwzględniając schemat z rys. 48, przypomnijmy poszczególne kroki dotyczące samego procesu implementacji WWW-ISIS:

- Na samym początku musimy przeanalizować strukturę bazy; zdecydować, które jej atrybuty (elementy) winny być przeszukiwane i na tej podstawie ustalić tablicę wyboru pól (FST) (etap 2.3: krok1, krok2).
- Następnie na podstawie ustalonej tablicy wyboru pól (FST) określamy tablicę definicji wyszukiwań (czyli po kolei ustalamy zawartość zbiorów: database.def..., browse.def...) (krok3, etap 2.4).
- Definiujemy niektóre spośród formatów dla naszej bazy (krok4, etap 2.5).
- Tworzymy zbiory HTML, określając właściwe połączenia hipertekstowe do interfejsów wyszukiwawczych (etap 2.6).
- Dokonujemy stosownych zmian w zbiorze konfiguracyjnym, a także w zbiorach HTTP (etap 2.7).
- Na koniec dokonujemy testowania systemu, a w przypadku konieczności powtarzamy poprzednie czynności (etap 2.8).

Przejdźmy zatem do szczegółowego omówienia poszczególnych etapów.

4.5.1. Analizowanie struktury bazy ISIS oraz tablicy selekcji pól – FST

Po pierwsze należy zauważyć, że nie wszystkie zbiory ISIS pełnią równie istotną rolę w systemie WWW-ISIS. Zbiorami nieistotnymi są formularze wprowadzania danych w wersji DOS oraz Windows. Podobnie zbiór FDT (tablica definicji pól) nie pełni decydującej roli w systemie WWW-ISIS, a korzystamy z niego jedynie po to, aby dowiedzieć się, jakiego rodzaju pola występują w naszej bazie.

Analiza struktury bazy ISIS

Rozpocznijmy proces tworzenia interfejsu WWW-ISIS od analizy struktury bazy ISIS, z punktu widzenia przyszłego procesu wyszukiwawczego. Projektant systemu winien odpowiedzieć na następujące pytania:

PYTANIE 1:

Czy zamierzamy dokonać rozróżnienia między różnymi typami rekordów?

☐

TAK

☐

NIE

PYTANIE 2:

Które z pól i w jaki sposób powinny być przeszukiwane w tworzonym interfejsie?

☐

TAK

☐

NIE

Odpowiedzi na obydwa pytania powinny być motywowane jedynie przyszłymi potrzebami użytkowników systemu wyszukiwawczego WWW-ISIS. Jeśli w naszym zamierzeniu użytkownicy będą chcieli dokonać rozróżnienia między dwoma typami rekordów bazy bibliograficznej: powiedzmy artykuły oraz książki, to winniśmy stworzyć dwa interfejsy wyszukiwawcze. Projektowanie wielu interfejsów wyszukiwawczych dla jednej bazy nie musi się bynajmniej ograniczać do różnych typów rekordów, a może dotyczyć wyświetlania tych samych rekordów w różny sposób.

Ustaliwszy typy przeszukiwanych pól wybieramy te pola, które będą zaindeksowane. Będąc bardziej precyzyjni możemy postrzegać indeksowanie bardziej ogólnie niż tylko jako bezpośrednie nawiązanie do tablicy definicji pól. Pamiętajmy, że ważniejsze w procesie tworzenia interfejsu WWW-ISIS są potrzeby informacyjne przyszłych użytkowników niż stopień szczegółowości czy poprawność bibliograficznego opisu danych.



UWAGA:

W przyszłym interfejsie wyszukiwawczym bierzemy pod uwagę tylko te pola, których znaczenie jest w pełni oczywiste dla użytkowników bazy. Powinniśmy unikać wprowadzania do interfejsu pól o niejasnym znaczeniu, zrozumiałym wyłącznie dla wąskiej grupy użytkowników.



ZAPAMIĘTAJ

Przedrostek „T=” w zbiorze FST odnosi się do typów rekordów.

4.5.1.1. Przykład bazy bibliograficznej AGRIN

Rozważmy jako przykład bazę AGRIN. Najpierw przyjrzyjmy się niezminionej tablicy wyboru pól.

```

2 0 v2
8 0 v8
9 1 (v9/)
1100 2 mp1,v1100
1110 0 v1110
1110 2 mp1,v1110
1200 4 mh1,v1200
2100 2 mp1,v2100
2110 0 v2110
2110 2 mp1,v2110
2200 4 mh1,v2200
4230 4 mh1,v4230
5800 2 mpu,v5800,v5801
5820 2 mpu,v5820
7400 0 v7400

```

Jak widać istniejące FST jest dość proste i nie może być bezpośrednio wykorzystane w tworzonym interfejsie WWW-ISIS głównie dlatego, że poszczególne indeksy nie są w nim rozróżnialne. Analizując FDT (tablicę wyboru pól) możemy zauważyć dwa typy rekordów wykorzystując następujące rodzaje bibliograficznego opisu:

Książki – (M, MS)

Artykuły – (AM, AMS, AS)

Jeśli założymy, że powyższe rozróżnienie może być istotne dla końcowego użytkownika, to konieczne jest przeprojektowanie zbioru FST (tablicy wyboru pól) tak, aby nowe FST odpowiadało przyjętym założeniom. W przeciwnym przypadku uznajemy wszystkie rekordy w bazie za opisy bibliograficzne tego samego typu.



ZAŁOŻENIE

Dla celów niniejszego przykładu zakładamy, że zamierzamy przygotować dla bazy AGRIN dwa interfejsy wyszukiwawcze: uproszczony (Wyszukiwanie uproszczone) oraz zaawansowany (Wyszukiwanie zaawansowane). Postanowiliśmy nie dokonywać rozróżnienia między typami rekordów traktując je jednakowo.

Wyszukiwanie uproszczone

Interfejs wyszukiwania uproszczonego powinien wyświetlać jedno pole (w formularzu HTML), do którego przyszli użytkownicy mogą wprowadzić określone terminy wyszukiwawcze. Terminy te powinny być wprowadzane na podstawie pól (bazy CDS/ISIS): wszystkich pól dotyczących, np. autora tekstu, tytułu (w tym tytuł czasopisma, jeśli występuje), abstraktu, deskryptorów, słów kluczowych wyodrębnionych z nazwy konferencji i lokalizacji konferencji.

Wyszukiwanie zaawansowane

Interfejs wyszukiwania zaawansowanego powinien wyświetlać wiele pól (w formularzu HTML), związanych z następującymi polami (bazy CDS/ISIS), jak np.:

- Słowa kluczowe z tytułu.
- Autor zbiorowy.
- Słowa kluczowe z nazwy konferencji (w tym lokalizacja konferencji).
- Deskryptory (wszelkie typy deskryptorów w tym lokalne).

Poniżej zostaną pokazane w postaci tabeli przyporządkowania odpowiednich przedrostków poszczególnym „indeksom” oraz sposób ich skonstruowania⁶.

Tabela 6
Tworzenie indeksów dla wyszukiwania uproszczonego (baza AGRIN)

Wirtualna nazwa indeksu	Istniejące pola w bazie ISIS		Uwagi	Proponowany prefiks
	Nazwa pola	Nr pola		
Wyszukiwanie uproszczone				
Keywords from record [słowa kluczowe z rekordu]	English Title monogr. Level	2200	Wybór przez słowa kluczowe	K=
	Original Title monogr. Level	2230	j.w.	
	English Title analytic level	1200	j.w.	
	Original Title analytic level	1230	j.w.	
	Serial main Title	4230	j.w.	
	Serial sec. Title	4231	j.w.	
	Author - monogr level	2100	Wybór słów ze znaczników < .>	
	Author - analytic level	1100	Wybór słów ze znaczników < .>	
	Corporate Author - monogr level	2110	Wybór przez słowa kluczowe	
	Corporate Author - analytic level	1110	Wybór przez słowa kluczowe	
	Conference title (analytic)	1210	Wybór przez słowa kluczowe	
	Conference title (monogr)	2210	Wybór przez słowa kluczowe	
	Conference place (analytic)	1211	Wybór z pól	
	Conference place (monogr)	2211	Wybór z pól	
	Agrovoc descriptors	5800	Wybór z zagnieżdzo- nych deskryptorów	
Proposed descriptors	5801	j.w.		
Local descriptors	5820	j.w.		

⁶ Tablica ta została opracowana przy założeniu, że językiem wiodącym dla bazy AGRIN jest język angielski (FDT pochodzi z 1992 r.).

Analogiczna tablica winna być sporządzona dla interfejsu zaawansowanego.

Tabela 7
Tworzenie indeksów dla wyszukiwania zaawansowanego (baza AGRIN)

Wirtualna nazwa indeksu	Istniejące pola w bazie ISIS		Uwagi	Proponowany prefiks
	Nazwa Pola	Nr pola		
Wyszukiwanie zaawansowane				
Słowa kluczowe z tytułu	English Title monogr. level	2200		TI=
	Original Title monogr. level	2230		
	English Title analytic level	1200		
	Original Title analytic level	1230		
	Serial main Title	4230		
	Serial sec. Title	4231		
Autor	Author - monogr level	2100	Dla artykułów pobieranie z pola 1100 ; dla książek pobieranie z pola 2100	A=
	Author - analytic level	1100		
Corporate Author	Corp. Author - monogr. Level	2110	Indeks dla słów kluczowych: dla artykułów pobieranie z pola 1100 ; dla książek pobieranie z pola 2100	C=
	Corp. Author - analytic Level	1110		
Keywords from Conference title	Conference title (analytic)	1210	Indeks na podstawie słów kluczowych	CO=
	Conference title (monogr)	2210		
	Conference place (analytic)	1211		
	Conference place (monogr)	2211		
Descriptors	Agrovoc descriptors	5800	Indeks na podstawie zagnieżdżonych słów kluczowych	D=
	Proposed descriptors	5801		
	Local descriptors	5820		

Konstrukcja nowego FST

Mając obie tablice (tab. 6 i 7), możemy przystąpić do tworzenia tablicy wyboru pól dla naszej przykładowej bazy AGRIN.

Tabela 8
FST (WWW-ISIS) dla bazy AGRIN

Nowy FST	Komentarz
1	2
2200 8 'K=/' V2200	Słowa kluczowe z tytułu dot. opisów monografii (wersja angielska)
2230 8 'K=/' V2230	Słowa kluczowe z tytułu dot. opisów monografii (wersja oryginalne)
1200 8 'K=/' V1200	Słowa kluczowe z tytułu dot. opisów artykułów (wersja angielska) Keywords from Analytic Title (English)
1230 8 'K=/' V1230	Słowa kluczowe z tytułu dot. opisów artykułów (wersja oryginalne)
4230 8 'K=/' V4230	Słowa kluczowe z tytułu czasopism

1	2
4231 8 'K=/' V4231	Słowa kluczowe z tytułu czasopism (Dodatek)
2100 6 'K=/' V2100	Autor (wydawnictwa zwarte)
1100 6 'K=/' V1000	Autor (czasopisma)
2110 8 'K=/' V2110	Autor zbiorowy dla wydawnictw zwartych (słowa kluczowe)
1110 8 'K=/' V1110	Autor zbiorowy dla artykułów (słowa kluczowe)
1210 8 'K=/' V1210	Słowa kluczowe z nazw konferencji (artykuły)
2210 8 'K=/' V2210	Słowa kluczowe z nazw konferencji (wydawnictwa zwarte)
1211 0 "K="V1211	Miejsce (artykuły)
2211 0 "K="V2211	Miejże (artykuły)
5800 6 0 'K=/'V5800	Deskryptory wg tezauryśa Agrovoc
5801 6 0 'K=/'V5801	Sugerowane terminy
5820 6 0 'K=/'V5820	Lokalne deskryptory
1200 8 'TI=/' V1200," "V1230," "V2200," "V2230," "V4230," "V4231	Wszystkie typy tytułów są umieszczone w indeksie i poprzedzone prefiksem TI=
2100 6 'A=/' if V11:'A' then V1100 else V2100 fi	Jeśli mamy do czynienia z rekordem opisującym artykuł (analytic), to wybierany jest z niego tylko autor, w przeciwnym przypadku pobierany jest z rekordu do indeksu autor wydawnictwa zwartego
2110 8 'C=/' if V11:'A' then V1110 else V2110 fi	Jeśli mamy do czynienia z rekordem opisującym artykuł (analytic), to wybierany jest z niego tylko autor zbiorowy, w przeciwnym przypadku pobierany jest z rekordu do indeksu autor wydawnictwa zwartego.
1210 8 'CO=/' if V11:'A' then V1210," "V1211 else V2210," "V2211 fi	Tytuł konferencji i miejsce konferencji tworzą wirtualny indeks poprzedzony prefiksem CO= (zarówno dla artykułów jak i wydawnictw zwartych)
5800 6 0 'D=/'V5800," "v5801," "v5820	

Po zapisaniu zbioru FST należy uruchomić pełną aktualizację zbioru odwróconego dla części bazy AGRIN i sprawdzić, czy indeksy zostały poprawnie utworzone.

WSKAZÓWKA

Jeśli tworzymy jeden interfejs dla kilku baz danych, to winniśmy zadbać, aby analogiczne elementy we wszystkich bazach miały te same prefiksy.



Przykład powyższy (tab. 8) pokazuje nam w jaki sposób należy przekształcać zbiór FST. Wspomniane FST powinno być zgodne z tablicą definicji pól, głównie ze względu na sposób identyfikacji rodzajów rekordów oraz przyporządkowania prefiksów do przeszukiwanych pól.

4.5.2. Tablice w WWW-ISIS

Jak już zostało wcześniej wspomniane, kluczową rolę w interfejsie wyszukiwawczym WWW-ISIS pełnią specjalne tablice definiujące zarówno proces wyszukiwania, jak i przeglądania danych. Nazwy i położenia owych tablic definicji WWW-ISIS określone są w zbiorze konfiguracyjnym isis3w.exe.cfg

Tablice definicji są zwykłymi zbiorami tekstowymi. Nazewnictwo zbiorów określone jest przez odpowiednią konwencję. Dla procesu wyszukiwania danych. nazwy zbiorów zaczynają się database.def. Zaś dla przeglądania danych nazwy

zbiorów zaczynają się browse.def. Ponadto dodawana jest końcówka określająca język etniczny interfejsu np.:

- Database.def.fr dla standardowych interfejsów wyszukiwawczych w języku francuskim.
- Database.def.sp dla standardowych interfejsów wyszukiwawczych w języku hiszpańskim itd.

Oczywiście, oprócz zdefiniowania powyższych tablic, niezbędne jest umieszczenie odpowiednich połączeń z zbiorach HTML (chodzi głównie o odwołania do skryptu cgi oraz o zmienną lang w połączeniach hipertekstowych). Jeśli o tym zapomnimy, system nie będzie działał poprawnie.

4.5.2.1. Tablica definicji wyszukiwań WWW-ISIS – ogólne zasady

Zanim omówimy szczegóły formalne najważniejszego elementu systemu wyszukiwawczego, przyjrzyjmy się dokładniej zbiorowi database.def.eng dla aplikacji DEMO (tab. 9).

Tabela 9

Omówienie tablicy wyszukiwań WWW-ISIS dla systemu DEMO

1.	# DEMO database def file created 8/2/1999		
2.	#Sekcja 1 - definicja wyszukiwania uproszczonego		Sekcja dla interfejsu uproszczonego
3.	[mons; PUB; Simple search]		
4.	[sort; @SCHR; Chronologically]	Formaty sortowania	
5.	[sort; @SAUT; by Author]		
6.	[sort; @STIT; by Title]		
7.	[format; @ESMON; Short format]	Format wyświetlania	
8.	[format; @EBMON; Bibliography format]		
9.	[format; @ELMON; Long format]		
10.	[format; @EFMON; Full format]	Opis przeszukiwanych pól	
11.	"Keywords from record", "keywords_record", "K=", "1";		
12.	# Sekcja 2 - definicja wyszukiwania zaawansowanego		
13.	[mona; PUB; Advanced search]		Sekcja dla interfejsu zaawansowanego
14.	[sort; @SCHR; Chronologically]	Formaty sortowania	
15.	[sort; @SAUT; by Author]		
16.	[sort; @STIT; by Title]		
17.	[format; @ESMON; Short format]	Format wyświetlania	
18.	[format; @EBMON; Bibliography]		
19.	[format; @ELMON; Long format]		
20.	[format; @EFMON; Full format]	Opis przeszukiwanych pól	
21.	"Keywords from Title", "mtitle", "TI=", "1";		
22.	"Personal Author", "mauthor", "A=", "1" ;		
23.	"Corporate Author", "mcorp author", "C=", "1";		
24.	"Keywords from Conference", "mconf", "CO=", "1";		
25.	"Descriptors", "mthesaurus", "D=", "1";		
26.	"Geographic Coverage", "mgeo", "GE=", "1";		
27.	"Language", "mlanguage", "L=", "1" ;		
28.	"Document type", "mdoc type", "DT=", "1";		
29.	"Year, e.g. 1981;1982", "myear", "YY=", "0";		
30.	# Sekcja 3 - Lista pól z ograniczoną liczbą znaków wyszukiwawczych		Lista pól z ograniczoną liczbą znaków wyszukiwawczych
31.	"restricted_terms", "mtitle", "mauthor", "mcorp_author", "mconf", "myear", "mgeo", "mthesaurus"		

Jak widać, w naszym zbiorze definicji WWW-ISIS wyróżnione są dwie sekcje (tablice):

- **mona** (sekcja 1, wiersze 2-11) dla wyszukiwania uproszczonego,
- **mons** (sekcja 2; wiersze 12-29) dla wyszukiwania zaawansowanego.

Z punktu widzenia samego systemu WWW-ISIS obie tablice zdefiniowane są dla tego samego typu rekordu PUB, co oznacza, że postrzegamy naszą bazę jako zbiór rekordów jednego typu (choć w samej bazie możemy dokonać rozróżnienia między artykułem a książką).

Każda tablica (przedstawiona w tab. 9) składa się z następujących specyfikacji:

- trzech formatów sortowania;
- czterech formatów wyświetlania danych;
- specyfikacji przeszukiwanych pól.

Sekcja dotycząca wyszukiwania uproszczonego definiuje tylko jedno pole określone jako „keywords from record” (słowa kluczowe z rekordu), podczas gdy sekcja dla wyszukiwania zaawansowanego odwołuje się do 9 pól. Znaczenie ostatniej sekcji #3 (wiersz 30 do 31) zostanie wyjaśnione w dalszej części.

Teraz możemy rozpocząć szczegółowe omawianie składni tablicy definicji wyszukiwań WWW-ISIS. Tablice definicji wyszukiwań podzielone są na osobne sekcje zwane tablicami (jedna tablica dla każdego typu rekordu, typy rekordów mogą jednak należeć do różnych baz ISIS, np.: wiersze 2-11, 13-29). Tablica składa się z następujących komponentów:

- wiersz opisujący typ rekordu;
- opcjonalne określenie zasad sortowania i formatowania;
- kolejności występowania zdefiniowanych elementów.

Ostatni komponent tablicy definicji WWW-ISIS określa listę pól z ograniczoną liczbą znaków wyszukiwawczych. Oprócz tego w dowolnym wierszu można umieszczać komentarze poprzedzone znakiem „#”.

4.5.2.2. Definicje typu rekordu

Wiersz opisujący typ rekordu (np. wiersze 3, 13) umieszczony jest w nawiasach kwadratowych i posiada następującą strukturę:

[nazwa_tablicy; ISISowa_nazwa_typu_danych; tekst_na_formularzu]
gdzie:

nazwa_tablicy jest nazwą typu rekordu używaną wewnętrznie przez program (skrypt) cgi; jeśli w danym połączeniu hipertekstowym chcemy umieścić odwołanie do danej tablicy, to zapis będzie miał mniej więcej następującą postać: „http://...&table=nazwa_tablicy&...”.

ISIS-owa_nazwa_typu_danych jest nazwą typu rekordu tak, jak jest ona używana w zapytaniach. Jeśli parametr IsisForceType w zbiorze konfiguracyjnym (isis3w.exe.cfg) jest równy 1, to program isis3w łączy domyślnie terminy operatorem AND.

T=type

Np. jeśli w przykładowej bazie nasze zapytanie wyszukiwawcze wygląda następująco: A=Brown +A= Smith

To system uruchomi następujące zapytanie: T=PUB*(A=Brown +A= Smith)

tekst_na_formularzu – jest nazwą, która jest wyświetlana użytkownikowi w formularzu wyszukiwawczym.

Przyjrzyjmy się jeszcze raz zapisowi z tab. 9: [mons; PUB; Simple search]
Pierwszy element definiuje tablicę mons dla typu rekordu ISIS, określonego jako PUB. W momencie wywołania tej tablicy program wyświetli informację: „Simple search”⁷.

4.5.2.3. Definicje wyboru pól

Definicja każdego z elementów pól określona jest przez cztery składniki:

- nazwę pola wyświetlaną w odpowiednim formularzu HTML,
- nazwę pola używaną wewnętrznie przez formularz HTML,
- prefiks ISIS wykorzystywany do zaindeksowania w zbiorze odwróconym,
- typ pola.

Liczba typów rekordów w zbiorze definicji pól WWW-ISIS oraz liczba elementów przypadająca na każdy typ rekordu są ograniczone. Składnia wygląda następująco:

```
"nazwa_pola_widoczna_na_formularzu","nazwa_pola_w_zbiorze_html",  
"Przedrostek_indeksujący_w_bazieISIS","typ_pola ", {"rozmiar_okna",  
"wartości_domyślne", "liczba_stałych_opcji", "opcja1", "opcja2"...};
```

W obrębie wiersza opisującego pole niedopuszczalne jest wprowadzanie pustych wierszy. Terminy w nawiasach klamrowych { } są opcjonalne. Można je używać jedynie w przypadku określenia typu pola jako 5 lub 6 (pole typu kombi).

Można zdefiniować następujące typy pól:

0 – zwykłe pole tekstowe bez żadnej listy słów kluczowych (nie pojawi się przycisk List);

1 – zwykłe pole tekstowe skojarzone z listą słów kluczowych (pojawia się przycisk List);

3 – statyczne pole tekstowe może być wykorzystane jako podtytuł dla całej grupy przeszukiwanych pól;

5 – statyczna lista z możliwością pojedynczego wyboru (lista możliwych do wyboru wartości jest dostępna przez pole typu kombi);

6 – statyczna lista z możliwością wielokrotnego wyboru (lista możliwych do wyboru wartości jest dostępna przez pole typu kombi).

Przykład 1

Poniższy wiersz z przykładowego systemu DEMO jest częścią tablicy definicji pól WWW-ISIS (tab. 5):

```
"Keywords from Title","mtitle", "TI=", "1";
```

Pierwszy składnik „Keywords from Title” jest tekstem, który zostanie wyświetlony w formularzu wyszukiwawczym HTML. Dla programu cgi pole to oznaczone jest jako mtitle, a słowa kluczowe będą używane w tworzonej formularzu HTML; zapis „TI=” jest prefiksem, który będzie wykorzystywany przez bazę ISIS; typ pola jest określony jako „1”. Tak więc, pole będzie wyświetlone z przyporządkowanym mu przyciskiem LIST (czyli ilekroć przycisk LIST będzie przyciśnięty, system będzie wyświetlał część zbioru odwróconego poprzedniego przedrostkiem „TI=”).

⁷ Pod warunkiem, że w zbiorze search.http (dla danej wersji językowej) nie jest zablokowany standardowy wiersz 10 (#10 %s%s).

Przykład 2



Założmy, że dla danego pola istnieje ograniczony zbiór dopuszczalnych wartości. W takim przypadku zamiast odwoływać się do zbioru odwróconego w celu znalezienia odpowiedniej wartości, możemy zdefiniować pole typu kombi z góry określonymi wartościami tak, aby końcowy użytkownik mógł wybrać stosowne wartości w celu utworzenia zapytania. Pole takie może być zdefiniowane przez określenie jego typu jako 5 lub 6. Poniższy zapis definiuje pole typu kombi tak, jak będzie ono wyglądało na rys. 49.

"Collection: ","docs", "FAO=", "5","1","0","4","", "FAO documents", "Library books", "Serials".

SEARCH EDIT QUERY UNDO CLEAR

Exact match of words: ☒

Sort: Chronologically

Format: Short format

Collection: FAO documents

FAO documents NOT

Library books

Serials

Full text

Rys. 49. Pole typu kombi w interfejsie wyszukiwawczym

W zapisie powyższym zdefiniowaliśmy pole Collection jako pole typu „kombi” pojedynczego wyboru (type=5) rozmiar okna jest równy 1, wartość domyślna jest równa (0), a lista opcji do wyboru składa się z 4 elementów, z których pierwszy jest pusty.

WSKAZÓWKA 1

Zaleca się określanie typu pola jako 0, jedynie wtedy, gdy nie jest wymagane odwoływanie się do zbioru odwróconego, np. rok, data itd.

WSKAZÓWKA 2

Zaleca się określanie typu pola jako 5 dla pól, których lista wartości nie jest zbyt obszerna, np. nazwa biblioteki, język dokumentu (E, F lub S) itd., zaś dopuszczalna liczba wartości nie jest większa niż 1.

WSKAZÓWKA 3

Zaleca się określanie typu pola jako 6 dla pól, których lista wartości nie jest zbyt obszerna, np. nazwa biblioteki, język dokumentu (E, F lub S) itd., zaś dopuszczalna liczba wartości przekracza 1.

Przyjrzyjmy się dokładniej specyfikacjom dotyczącym sortowania wyników oraz wyświetlanych formatów.

4.5.2.4. Parametry sortowania wyników

W projektowaniu interfejsu wyszukiwawczego WWW-ISIS może okazać się konieczne zdefiniowanie pól typu kombi pozwalających na posortowanie poszukiwanych rekordów.

Specyfikacje dotyczące sortowania pojawiają się zaraz po definicjach typu rekordu w następujący sposób:

```
[sort; format_sortowania_w_isis1; tekst_na_formularzu1]  
[sort; format_sortowania_w_isis2; tekst_na_formularzu2]  
.....  
[sort; format_sortowania_w_isisn; tekst_na_formularzun]
```

gdzie:

sort jest słowem kluczowym identyfikującym opcje formatu;

format_sortowania_w_isis jest nazwą formatu wyrażoną w języku formatowania (używanym do posortowania danych);

tekst_na_formularzu jest nazwą wyświetlaną w formularzu HTML (w rozwijanym menu pola typu kombi).

Przykład 3

Przykłady specyfikacji sortowania dla systemu DEMO są pokazane poniżej:

```
[sort; @SCHR; Chronologically]  
[sort; @SAUT; by Author]  
[sort; @STIT; by Title]
```

Dzięki powyższemu zapisowi projektant może uzyskać następujący ekran rys. 50:

The screenshot shows a search interface with the following elements:

- Exact match of words:** A checkbox that is checked.
- Sort:** A dropdown menu currently showing "Chronologically". A callout box points to this menu with the text: "Pole typu kombi do wyboru sposobu sortowania znalezionych wyników".
- Format:** A dropdown menu currently showing "Short format". A callout box points to this menu with the text: "Pole typu kombi do wyboru formatów wyświetlania znalezionych wyników".
- Keywords from Record:** A text input field.
- OR:** A dropdown menu.
- List:** A button.

Rys. 50. Zastosowanie formatów wyświetlania i sortowania

Naturalnie w przypadku projektowania pól do sortowania znalezionych danych należy również przygotować odpowiednie formaty ISIS tworzące klucze sortowania danych.

Przykład 4

Przygotujmy format, który tworzy odpowiednie klucze sortowania zgodnie z polami typu autor. Jeśli pojedynczy autor nie istnieje, to powinien być uwzględniony autor zbiorowy. Zakładamy, że projektant będzie pracował z bazą DEMO, w której informacje nt. pojedynczego autora znajdują się w polu oznaczonym znacznikiem V28^a (pole nr 28, podpole a), zaś tytuł wydawnictwa zwartego umieszczony jest zwykle w polu oznaczonym znacznikiem V29^a (pole o znaczniku 29, podpole a) zaś autor zbiorowy odpowiednio w polach 30 (V30) oraz 31 (V31).



W takim przypadku format powinien wyglądać następująco:

```
MHU,  
if p(v28) then V28^a.35 else  
  if p(v30) then V30.35 else  
    if p(v29) then V29^a.35 else  
      if p(v31) then V31.35,  
    fi fi fi fi
```

Zauważmy, że klucz sortujący jest tworzony w ten sposób, że są pobierane wszystkie wystąpienia pola Autor (Author), następnie uzyskany ciąg jest ograniczany do 35 znaków. W rzeczywistości oznacza to, że rekordy będą sortowane według wystąpienia tylko pierwszego autora.

Przykład 5

Poniżej pokażemy jak zaprojektować klucze sortujące w porządku odwrotnym do chronologicznego. Zakładamy, że pole V64 przechowuje w pierwszych 4 znakach informacje nt. roku publikacji. Dlatego też klucz sortujący powinien wyglądać następująco:



$s(f(3000 - val(v64.4), 4, 0))$,

Oznacza to, że najpierw będą wyświetlane najnowsze publikacje.

4.5.2.5. Parametry określające format wyświetlanych wyników

Specyfikacje dotyczące formatu wyświetlanych danych pozwalają użytkownikowi na wybranie określonego formatu wyświetlania danych przez skorzystanie z listy rozwijanych menu (w polu typu kombi, patrz rys. 51).



Format:

Keywords from Record

Short format

Short format

Bibliography format

Long format

Full format

Rys. 51. Przykład pola kombi do wyboru formatów wyświetlania danych

Instrukcje dotyczące formatowania mają następującą składnię:

```
[format; nazwa_zbioru_formatującego_wISIS1; tytuł_formatu_na_formularzu1]  
[format; nazwa_zbioru_formatującego_wISIS2; tytuł_formatu_na_formularzu2]  
...  
[format; nazwa_zbioru_formatującego_wISISn; tytuł_formatu_na_formularzun]
```

gdzie:

format jest słowem kluczowym identyfikującym opcje formatu,
nazwa_zbioru_formatującego_wISIS jest nazwą formatu w ISIS,
tytuł_formatu_na_formularzu jest tekstem wyświetlanym w formularzu HTML w polu typu kombi.

Dla bazy DEMO został zdefiniowany następujący format (którego wynik pokazany jest na rys. 51):

```
[format; @ESMON; Short format]  
[format; @EBMON; Bibliography format]  
[format; @ELMON; Long format]  
[format; @EFMON; Full format]
```



UWAGA:

Domyślnie pierwszy format jest wartością wyświetlaną w polu typu kombi (rys. 50). Jeśli lista zawiera tylko jeden element, to nie zostanie wyświetlone pole typu kombi, chociaż odpowiedź systemu będzie wyświetlona przy użyciu znaczników HTML z atrybutem HIDDEN.



4.5.2.6. Lista pól z ograniczoną liczbą znaków wyszukiwawczych

Ostatnia sekcja tablicy wyszukiwania pól WWW-ISIS to lista pól, dla których system sprawdza, czy terminy nie są zbyt krótkie – chodzi tu o niedopuszczenie do sytuacji blokowania serwera pytaniami, które wymagają długich czasów obsługi. Mimo że sekcja ta nie jest obowiązkowa, to zaleca się jej wprowadzenie. Składnia wygląda następująco:

"restricted_terms", "termin₁", ... " termin_n"

Lista pól z ograniczoną liczbą znaków wyszukiwawczych zaczyna się od podania słowa kluczowego, a następnie nazw pól (w cudzysłowach). Lista ta określa, w których polach wyrażenia wyszukiwawcze nie mogą być krótsze niż wartość podana w zbiorze konfiguracyjnym (patrz parametr RestrictedTermMinLength" w zbiorze isis3w.exe.cfg, Aneks A). Ograniczenie to jest spowodowane specyfiką procesu wyszukiwawczego z zastosowaniem znaków globalnych (w ISIS jest to \$). Jeśli termin wyszukiwawczy jest zbyt krótki (np. wprowadzamy a i nie zaznaczamy pola „Exact match”, czyli żądamy znalezienia wszystkich terminów zaczynających się na a), to takie zlecenie może zablokować serwer. W języku wyszukiwawczym ISIS zapytanie takie miałoby postać:

TYT=a\$

Aby uniknąć niepotrzebnego przeciążenia systemu wprowadza się sekcje restricted_terms wraz z odpowiednimi ustawieniami w zbiorze konfiguracyjnym.

UWAGA:

Zaleca się dodawanie do listy pól z ograniczoną liczbą znaków wyszukiwawczych, dla których indeksy są odpowiednio długie. Nigdy nie wprowadzaj do listy pól, które zawierają określoną liczbę krótkich oznaczeń kodowych (np. kod języka, kraju itd.)



4.5.2.7. Testowanie poprawności tablicy wyszukiwań

Po zapoznaniu się z tablicą FST oraz tablicą definicji wyszukiwań WWW-ISIS czytelnik powinien być w stanie przygotować odpowiedni zbiór database.def.eng. W celu przeprowadzenia testu należy:

- upewnić się, czy baza jest umieszczona w odpowiednim katalogu, określonym przez parametr IsisDataDir (lub w pliku określonym przez parametr DBNFile) w zbiorze isis3w.exe.cfg. Na tym poziomie nie jest konieczna znajomość wszystkich parametrów konfiguracyjnych zbioru isis3w.exe.cfg⁸;
- z bazy DEMO skopiuj format ephmon.pft;
- upewnij się, czy nowy zbiór FST jest w tym samym katalogu, a zbiór odwrócony został uaktualniony;
- w katalogu macierzystym serwera WWW utwórz katalog dla twojej aplikacji WWW-ISIS, a następnie skopiuj tam tablicę definicji pól WWW-ISIS – database.def.eng;
- skopiuj z katalogu DEMO do nowo utworzonego katalogu zbiory HTML oraz podkatalogi HELP, HTP oraz IMGS.

Następnie w katalogu cgi-bin (skryptów serwera WWW) zmień nazwę zbioru isis3w.exe.cfg na dowolną inną. Skopiuj, zapisz i otwórz w dowolnym edytorze tekstowym zbiór isis3wdemo.exe.cfg i zmień w nim odwołania do tablic definicji pól WWW-ISIS oraz zbiorów HTP:

HtpFilePath	c:\waicent1\WEB\demo\htp
IncludeFilePath	c:\waicent1\WEB\demo\htp\eng
SearchDefinitionFile	c:\waicent1\WEB\demo\database.def
EntryDefinitionFile	c:\waicent1\WEB\demo\entry.def
BrowseDefinitionFile	c:\waicent1\WEB\demo\browse.def

UWAGA:

Uwaga: ciąg „c:\waicent1\WEB\demo” należy zastąpić właściwą lokalizacją zbiorów.



W zbiorze index.html znajdź odpowiedni wiersz wskazujący na odwołania do programu cgi (poszukując ciąg znaków cgi-bin); Wiersz ten powinien mieć postać:

```
<FRAME SRC="/cgi-bin/isis3w.exe?database=demo&search_type=entry&lang=eng&table=mons" NAME="display">
```

⁸ Pełny opis zbioru konfiguracyjnego znajduje się w Aneksie A.

Zmień w nim wartości w miejscach wytłuszczenia:

...database=demo...na ...database="nazwa nowo tworzonej bazy"...

...table=mons... na table="nazwa nowo utworzonej tablicy wyszukiwań"

czyli wprowadź nazwę swojej bazy danych oraz odpowiednie nazwy z tablicy definicji wyszukiwań WWW-ISIS (zbiór database.def.).

Dokonaj analogicznych zmian w zbiorze HTML toc.html w szczególności dokonaj modyfikacji połączeń hipermedialnych przycisków Simple Search (*Wyszukiwanie Uprozczone*) oraz Advanced Search (*Wyszukiwanie Zaawansowane*) zmieniając odpowiednie parametry database oraz table.

Następnie należy sprawdzić, czy przycisk LIST wyświetla właściwą część indeksu. Jeśli nie, należy sprawdzić, czy FST jest zgodny ze zbiorem database.def.

Przy założeniu, że wszystkie kroki zostaną wykonane z odpowiednią starannością, system powinien działać. Jednak pozostaje do omówienia kilka istotnych aspektów, w szczególności kwestia formatów wyświetlania danych.

4.5.3. Formaty⁹

4.5.3.1. Formaty wyświetlania rekordów

Z formalnego punktu widzenia formaty CDS/ISIS przygotowane dla wersji DOS lub Windows mogą być z powodzeniem wykorzystane w środowisku WWW-ISIS bez specjalnych modyfikacji. Jednak takie rozwiązanie nie zawsze wychodzi naprzeciw oczekiwaniom przyszłych użytkowników interfejsu WWW-ISIS i nie wykorzystuje w pełni wszystkich możliwości systemu. Zaleca się dokonanie pewnych modyfikacji formatów, przynajmniej dodawanie znaczników końca wiersza języka HTML. Jednak najlepszym rozwiązaniem jest stworzenie nowych formatów, a przynajmniej modyfikacja istniejących.

Dla każdego rekordu określonego w tablicy definicji wyszukiwań WWW-ISIS należy przygotować właściwy format przewidując odpowiedni wybór stosownych pól typu kombi. Formatom winny być przyporządkowane odpowiednie nazwy, ustalone w tablicy definicji wyszukiwań WWW-ISIS (por. 4.5.2.4).

Należy jednak pamiętać, że lista wyboru ustalona w formacie nie jest obligatoryjna. Jeśli nie pojawi się żadna lista wyboru, system przyjmie format domyślny. Nazwa domyślnego formatu jest ustalona w zbiorze isis3w.exe.cfg (zmienna DefaultFormat).

Zwykle dla każdego typu rekordu powinno zaprojektować się co najmniej dwa formaty wyświetlania danych: pełny i skrócony. Format skrócony pozwala na wyświetlanie jedynie najważniejszych danych. Na przykład w przypadku baz bibliograficznych można wyświetlić dane o autorze oraz tytule dokumentu, a także o dacie publikacji. Wyświetlone dane powinny także zawierać połączenie do pełnych danych rekordu.

Aby ułatwić administrowanie systemem znacznie wygodniej jest stosować odpowiednią konwencję w nazewnictwie formatów tak, aby bez większych problemów można było rozpoznać odpowiedni format. W szczególności, kiedy

⁹ Podrozdział ten poświęcony jest czytelnikowi zaznajomionemu z językiem formatowania ISIS.

zamierzamy stworzyć interfejs wielojęzyczny, nazwy odpowiednich formatów powinny uwzględniać informację, do jakiego języka etnicznego należą.

Formaty wyświetlania danych przygotowywane są zgodnie ze standardowym sposobem przygotowywania formatów w ISIS. Dozwolone jest wykorzystywanie niemal wszystkich technik formatowania ISIS w celu wyświetlenia rekordu w standardzie HTML.

UWAGA:

Nie wolno stosować następujących poleceń formatowania CDS/ISIS: polecenia wyrównania tekstu poziomego i pionowego (Xn, Cn, /, # oraz %). Zamiast tego można stosować odpowiednie polecenia HTML jako zagnieżdżone literały w formacie ISIS (por. Aneks B).



Przykład 6

Przyjrzyjmy się bliżej formatowi ESMON dla bazy DEMO (zapisanym w pliku ESMON.PFT).



Tablica 3

Format ESMON.PFT dla bazy DEMO

1.	mhl,		
2.	'<dl><dt>'/		
3.	if v11.1='A' then		
4.	'<A href = "/cgi-bin/isis3w.exe?rec_id=',		
5.	mfn,		
6.	'&database=demo&search_type=link&table=mona&lang=eng&format_name=ELMON">'		
7.	v33^e,"; "v35^e,		
8.	if a(v33^e) then		
9.	v33^f,"; "v35^f,		
10.	if a(v33^f) then		
11.	v33^s,"; "v35^s,		
12.	if a(v33^s) then		
13.	v33^o,"; "v35^o		
14.	fi,		
15.	fi,		
16.	fi,		
17.	' '/		
18.	if a(v33^e) then " ["v33^t,"; "v35^t,"], "d33^t fi		
19.	(, +v28^a)/		
20.	if p(v34) or p(v41) then ' In: ',' v34^e,		

Dotyczy rekordów typu artykuł (PAPER) wiersze 3-63

Połączenie hipertekstowe z wyświetlanym rekordem



Wyświetlanie tytułów w języku angielskim, francuskim lub hiszpańskim

Wyświetla informacje o publikacjach zwartych, jeśli dany artykuł do nich należy ("paper from book"), lub informacje dotyczące czasopism, jeśli dany artykuł pochodzi z czasopisma ("paper from a journal").

21.	if a(v34^e) then	
22.	v34^f,	
23.	if a(v34^f) then v34^s,	
24.	if a(v34^s) then v34^o fi,	
25.	fi,	
26.	fi,	
27.	if a(v34^e) then " ["v34^t"]" fi/	
28.	if p(v34) and p(v41) then '; ' fi/	
29.	v41^e,	
30.	if a(v41^e) then v41^f,	
31.	if a(v41^f) then v41^s,	
32.	if a(v41^s) then v41	
33.	fi,	
34.	fi,	
35.	fi,/	
36.	fi,	
37.	if p(v46) then '<I>',	
38.	if p(v28^a) or p(v41) or p(v34) then '; ' fi,	
39.	v46^e,	
40.	if a(v46^e) then v46^f,	
41.	if a(v46^f) then v46^s,	
42.	if a(v46^s) then v46^o	
43.	fi,	
44.	fi,	
45.	fi,	
46.	'</I>',	
47.	fi,	
48.	", "v46^n,", "v48,", "v50/	
49.	(, v29^a, <I> v29^b </I>)/	
50.	if (p(v28^a) or p(v29^a) or p(v41) or p(v46)) and (p(v31) or p(v53)) then	
51.	' / '	
52.	fi/	
53.	v31^e+ ; ,	
54.	if a(v31^e) then v31^f+ ; ,	
55.	if a(v31^f) then v31^s+ ; ,	
56.	if a(v31^s) then v31^o+ ; fi,	
57.	fi,	
58.	fi,	
59.	if a(v31) then v53,	
60.	if p(v53) and p(v52) then ', 'fi,v52	
61.	fi/	
62.	", "v64/	

63.	"", "v56^a/	
64.	Else	
65.	'<A href = "/cgi-bin/isis3w.exe?rec_id=',	
66.	mfn,	
67.	'&database=demo&search_type=link&table=mona&lang=eng&format_name=ELMON">'	
68.	,v34^e,"; "v36^e,	
69.	if a(v34^e) then v34^f,"; "v36^f,	
70.	if a(v34^f) then v34^s,"; "v36^s,	
71.	if a(v34^s) then v34^o,"; "v36^o fi,	
72.	fi,	
73.	fi,	
74.	' '	
75.	if a(v34^e) then " ["v34^t,"; "v36^t,"]. "d34^t fi	
76.	(, +v29^a, <I> v29^b </I>)/	
77.	if p(v41) then ' In: '	
78.	v41^e,	
79.	if a(v41^e) then v41^f,	
80.	if a(v41^f) then v41^s,	
81.	if a(v41^s) then v41 fi,	
82.	fi,	
83.	fi,/	
84.	fi,	
85.	if p(v46) then '<I>'	
86.	if p(v29^a) or p(v41) then '; ' fi,	
87.	v46^e,	
88.	if a(v46^e) then v46^f,	
89.	if a(v46^f) then v46^s,	
90.	if a(v46^s) then v46^o fi,	
91.	fi,	
92.	fi,	
93.	'</I>'	
94.	fi,	
95.	"", "v46^n,", "v48,", "v50/	
96.	if (p(v29^a) or p(v41) or p(v46)) and (p(v31) or p(v53)) then ' / ' fi/	
97.	v31^e+ ; ,	
98.	if a(v31^e) then v31^f+ ; ,	
99.	if a(v31^f) then v31^s+ ; ,	
100.	if a(v31^s) then v31^o+ ; fi,	
101.	fi,	
102.	fi,	
103.	if a(v31) then v53,	

Dotyczy typu książka (BOOK) wiersze: 64-109

Połączenie hipertekstowe z wyświetlanym rekordem 	Wyświetlanie tytułów książek
---	---

104.	if p(v53) and p(v52) then ', 'fi,	
105.	v52	
106.	fi/	
107.	", "v54/	
108.	", "v56^a/	
109.	fi,	
110.	'</dl>'	

Wszystkie wiersze formatu z tablicy 3 zostały celowo ponumerowane tak, aby łatwiej było dokonać ich analizy. Wiersze 3-63 odpowiadają za wyświetlanie danych związanych z wyświetlaniem rekordów typu „artykuł” (PAPER), podczas gdy wiersze 64-109 odpowiadają za wyświetlanie rekordów typu „książka”.

Wiersze 4-6 oraz 65-73 odpowiedzialne są za stworzenie połączenia z wyświetlanym rekordem (search_type=link), dostarczając tym samym innego formatu (do wyświetlania tego samego rekordu). Tak więc użytkownik WWW-ISIS korzystając z formatu ESMON.pft może kliknąć na połączeniu w celu wyświetlenia pełnego rekordu (utworzonego na podstawie innego formatu ELMON.pft).

Połączenie jest widoczne jako tytuł – wiersze 7-17 wyświetlają tytuł w języku angielskim lub francuskim czy hiszpańskim), podczas gdy wiersze 68-73 odpowiedzialne są za wyświetlanie tytułu książki w rekordach typu „książka” (BOOK).

Wiersz 18 wyświetla informacje o publikacjach zwartych, jeśli dany artykuł do nich należy („paper from book”) lub wyświetla informacje dotyczące czasopism, jeśli dany artykuł pochodzi z czasopisma („paper from a journal”).

Warto zauważyć, że połączenia są tworzone dynamicznie – wykorzystywany jest numer rekordu MFN (Master File Number). Można więc w dość prosty sposób zaadaptować powyższą strukturę w innych formatach przez zmianę poniższych parametrów:



URL:



database – parametr odpowiedzialny za wybór bazy;

table – parametr odpowiedzialny za wybór tablicy w zbiorze definicji wyszukiwań WWW-ISIS (database.def.eng);

format_name – nazwa formatu, do którego odwołuje się dane połączenie, np w celu wyświetlenia bardziej szczegółowego opisu dokumentu.



UWAGA:

Zauważ, że znaczniki HTML w formacie (nowy akapit, wytłuszczenie, nowy wiersz, pismo pochyle) wprowadzane są jako literały warunkowe lub bezwarunkowe.

Standardowy format powinien być zdefiniowany w podobny sposób (przez dodanie większej liczby specyfikacji dotyczących wyprowadzania danych z pól oraz podpól rekordów). Czasami może okazać się pożyteczne stworzenie tabeli HTML z dwiema kolumnami: nazwa pola (w pierwszej kolumnie) oraz zawartość pola (w drugiej kolumnie).

Poniżej przedstawiony jest fragment takiego formatu dla bazy DEMO:

Tabela 10

Przykładowy format dla bazy DEMO (pole/zawartość)

```
Mhl
'<dl><dt><table CELLPADDING=0>'
...
"<TR><TD BGCOLOR='#E8E8CC'><B><I>Division:</I></B></TD><TD>"d19
(v19+|;|)"</TD></TR>"d19/,
"<TR><TD BGCOLOR='#E8E8CC'><B><I>Publisher:</I></B></TD><TD>"v53"</TD></TR>",
"<TR><TD BGCOLOR='#E8E8CC'><B><I>Publ.place:</I></B></TD><TD>"v52"</TD></TR>",
"<TR><TD BGCOLOR='#E8E8CC'><B><I>Publ.date:</I></B></TD><TD>"v54"</TD></TR>",
"<TR><TD BGCOLOR='#E8E8CC'><B><I>Edition:</I></B></TD><TD>"v51"</TD></TR>",
"<TR><TD BGCOLOR='#E8E8CC'><B><I>Collation:</I></B></TD><TD>"v56"</TD></TR>",
"<TR><TD BGCOLOR='#E8E8CC'><B><I>Collation(S):</I></B></TD><TD>"v57"</TD></TR>",
"<TR><TD BGCOLOR='#E8E8CC'><B><I>ISSN:</I></B></TD><TD>"v27"</TD></TR>",
"<TR><TD BGCOLOR='#E8E8CC'><B><I>ISBN:</I></B></TD><TD>"v26"</TD></TR>",
...
'</TABLE></dl>'
```

Podany w tab. 10 format, czytelnik może z łatwością zaadaptować do innej bazy, zmieniając nazwy pól oraz ich znaczniki.

Jak już zostało wcześniej wspomniane, język formatowania CDS/ISIS daje bardzo szerokie możliwości tworzenia różnego rodzaju połączeń. Są one szczególnie użyteczne w przypadku dynamicznych połączeń hipertekstowych w języku HTML. Poniżej opisane są najważniejsze z nich (patrz parametr `search_type`).

URL:

search_type=link – pozwala odwołanie się do numeru rekordu (MFN w bazie ISIS). W tym przypadku może być wzięty pod uwagę bieżący numer rekordu. Jeśli zamierzamy wyświetlić cały rekord (jest to standardowe działanie, gdy zamierzamy zmienić wyświetlanie danych z formatu skróconego na pełny)¹⁰.



URL:

search_type=query – daje możliwość wykorzystania wartości z pola lub z kilku pól odwołując się do tej samej lub innej bazy. Takie połączenie odwołuje się w zapytaniu do programu cgi, a końcowy użytkownik otrzymuje konkretną odpowiedź na zapytanie.



¹⁰ W przypadku języka formatowania ISIS, MFN może być sprawdzane przez funkcję REF, jednak funkcja ta wymaga udoskonalenia.



URL:



search_type=browse – jest wykorzystywane kiedy rekord dostarcza przedrostka (prefiksu), który jednocześnie służy jako przedrostek w wyszukiwaniu danych. Ta technika jest wykorzystywana w nawigacji w teaurusie (np. przeglądając terminy przechodzimy z porządku systematycznego do np. alfabetycznego).



URL:



search_type=add_thes_word – połączenia tego typu są wykorzystywane w przypadku wyszukiwania wykorzystującego teaurus. Z połączeniem tym skojarzony jest w takich przypadkach odpowiedni przycisk, który powoduje dodanie zaznaczonego wcześniej deskryptora do listy terminów wyszukiwanych.

Następujący przykład, wykorzystujący strukturę bazy CDS, pokazuje jak stworzyć połączenie, które pozwala na wyświetlanie wszystkich artykułów napisanych przez tego samego autora.



```
mhl, '<P>', '<B>', v24, '</B>', " / "d70,
(if p(v70) then
  '<A HREF="/cgi-bin/isis3w.exe?query=',
  'IA=', v70,
  '&database=cds&search_type=query&table=ana&lang=eng&format_name=@ESANA">',
  v70+|; |,, '</A>',
fi),
if p(v26^c) then '<BR>', '<I>', v26^c, '</I>' fi,
```

W powyższym formacie dynamiczne połączenie (search_type=query) jest skojarzone z polem autor (v70). Tak więc format wygeneruje następujący tekst dla każdego autora („Smith, S.T.; Brown, K.L.”) dołączając odpowiednie połączenia (pozwalające na dostęp do wszystkich jego prac):

Horticultural production in Zimbabwe / Smith, S.T.; Brown, K.L. The Uruguay Round and agriculture in Southern Africa. Implications and policy responses, Konandreas, P. (ed.), Lindland, J. (ed.), Pearce, R. (ed.), Wilkin, K.

W powyższym przykładzie każde połączenie odwołuje się do isis3w.exe w zapytaniach odnośnie innych artykułów, których autorami są podane osoby. Zapytanie będzie przesłane do serwera po „kliknięciu myszą” na nazwisku autora.

Aby właściwie zaprojektować odpowiednie połączenie należy odwołać się do struktury FST. W podanym tu przykładzie zakładamy, że dla tworzenia zapytań odnośnie konkretnego autora winniśmy zbudować odpowiednie zapytanie:

IA=name



UWAGA:

Ponieważ informacje o autorach są przechowywane w polach powtarzalnych, połączenie powinno być generowane wewnątrz powtarzalnej grupy. Szczególne występowanie tej grupy (tj. zapytań) powinno być rozdzielone średnikiem (; | +d70). Wskazany prefiks (IA=) musi być taki sam, jak zdefiniowany w FST dla terminów wyselekcjonowanych z pola 70. Wartość dla parametru „table=” musi być taka sama jak zdefiniowana w tablicy definicji wyszukiwań WWW-ISIS. W zależności od zapotrzebowania można dodać do formatu dowolnie wiele połączeń (np. search_type=query połączenia dla deskryptorów, nazw wydawnictw itd.). Nieco bardziej złożone przykłady przedstawionych typów połączeń oraz szczegółowy opis połączeń: search_type=browse oraz search_type=add_thes_word przedstawiony jest w dokumentacji technicznej „WWW/ISIS Technical Reference Manual version 5.0.0”.

4.5.3.2. Inne przydatne formaty

Obecnie przedyskutujemy inne typy formatów, których działanie w znacznym stopniu może poprawić funkcjonowanie systemu. W szczególności przedstawimy: format nagłówka strony [**page header format**], format stopki strony [**page footer format**], oraz nagłówka rekordu [**record header format**].

Głównym celem formatu nagłówka strony jest dostarczenie informacji na temat rezultatów uzyskanych przez isis3w. Czasami może okazać się użyteczne wyświetlenie informacji na temat liczby wyszukanych i wyświetlonych rekordów. Inną funkcją formatu nagłówka i stopki strony mogłaby być sytuacja, kiedy użytkownik systemu pragnie uzyskać dane w formie tabeli. W takim przypadku nagłówki może wprowadzić znaczniki HTML otwierające tabelę, zaś stopka może wprowadzić znaczniki HTML zamykające tabelę.

Budując format nagłówkowy można wykorzystywać specjalne parametry z tzw. wirtualnych pól:

V9999 – liczba rekordów odpowiadających ustalonym kryteriom wyszukiwawczym;

V9991 – liczba rekordów wybranych do wyświetlenia (ograniczenie jest ustalone na podstawie odpowiedniego parametru w zbiorze konfiguracyjnym – „QueryMaxRows”).

Poniżej przedstawiony jest przykład formatu nagłówkowego dla aplikacji DEMO:

```
'<CENTER><H4>',
if val(V9999) > val(V9991) then,
  V9999" record(s) found – ",
  V9991" record(s) presented (too general query)",
  '<BR> <A HREF=" ../demo/help/examples.html">See examples</A>'
else,
  if val(V9999) > 0 then,
    V9999" record(s) found",
  else,
    'Too restrictive query, ',
```



```
'<A HREF=" ../demo/help/examples.html">see examples</A>',  
fi,  
fi,  
'</H4></CENTER> <HR>',
```

Jak zostało wcześniej wspomniane, możliwe jest wykorzystanie formatu nagłówka rekordu. W takim przypadku wykorzystujemy wirtualne pole V9998, które przechowuje kolejne numery znalezionych rekordów.



URL:

Wszystkie wyżej wymienione formaty (nagłówkowe, strony, rekordu oraz stopki) są wywoływane przez parametry: **page_header**, **page_footer**, **record_header**.



4.5.4. Jak zaprojektować i przetestować formaty

Technika budowania formatów i ich testowania w środowisku WWW-ISIS jest w dużym stopniu odmienna od sposobu, w jaki formaty wykorzystuje się w środowisku ISIS. Poniżej znajduje się kilka rad ułatwiających zaprojektowanie i przetestowanie formatów środowisku WWW-ISIS:

Na początek radzilibyśmy rozpoczęcie pracy od zmiany istniejących i działających formatów CDS/ISIS (wersja DOS).

1. WPROWADŹ INSTRUKCJĘ „IF” NA POCZĄTKU I NA KOŃCU GRUPY POWTARZALNEJ.

CDS/ISIS
FORMATY



Konieczna byłaby taka zmiana ich struktury, która umożliwiłaby stosowanie instrukcji warunkowych (zagnieżdżonych wyrażeń „if”) na początku i końcu grupy powtarzalnej (tabl. 3 Format ESMON.PFT dla bazy DEMO).

2. ZASTĄP STANDARDOWE POLECENIA FORMATOWANIA TEKSTU W CDS/ISIS POLECENIAMI JĘZYKA HTML.

CDS/ISIS
FORMATY



Następnie należałoby usunąć z istniejących formatów typowe dla CDS/ISIS polecenia: wcięcie tekstu, wprowadzanie nowego wiersza, tabulacji itd. Dodaj do formatu odpowiednie (analogiczne) znaczniki HTML (jako literały).

3. PRZETESTUJ WYSZUKIWANIE W WWW-ISIS I/LUB DOS ISIS.



Budując format przeprowadź proste wyszukiwanie przy wykorzystaniu WWW-ISIS i przetestuj je online. Jeśli zostały popełnione jakieś błędy, to powinny być one widoczne w przeglądarce WWW. Jeśli nic nie jest wyświetlane, przetestuj format w systemie DOS ISIS.

Jeśli podstawowy format działa, można dodać konieczne połączenia. Sprawdź uważnie wszystkie połączenia. Jeśli nie działają wykorzystaj funkcje przeglądarki, aby skopiować połączenia, a następnie wklej je do dowolnego edytora tekstowego w celu szczegółowej analizy. Pamiętaj, że nie jest dozwolone wprowadzanie wewnątrz połączeń żadnych wolnych miejsc, podobnie nie należy wprowadzać żadnych specjalnych znaków, jak +, *, %. W przypadku wątpliwości radzimy szczegółowo zapoznać się z językiem HTML.

Zakładamy, że czytelnik działając zgodnie z przyjętym planem i wykonując krok po kroku opisane działania zdołał już zaimplementować system. Nadal jednak jest możliwość jego udoskonalania. W dalszej części przedyskutujemy, jak dokonać niezbędnych zmian w zbiorach HTML, HTTP oraz zbiorze konfiguracyjnym (isis3w.exe.cfg).

4.6. Modyfikacja zbiorów HTML oraz zbiorów HTTP

Czytelnik ma już za sobą pierwsze zmagania z konfiguracją zbioru „isis3w.exe.cfg”. Niniejszy podrozdział przedstawia szczegółowo zagadnienia konfiguracji systemu tak, aby interfejs w pełni odpowiadał oczekiwaniom przyszłych użytkowników.

4.6.1. Zbiory HTML

Przypomnijmy, że zbiory HTML są przechowywane w katalogu macierzystym serwera WWW – w przypadku XITAMI przy założeniu, że korzystamy z ustawień domyślnych, jest to C:\webpages\„Nazwa aplikacji”, a w przypadku aplikacji DEMO katalog ten to C:\webpages\demo.

Zbiorem początkowym jest index.html, który odwołuje się do zbioru toc.html. Zbiór toc.html jest wyświetlany w lewej ramce. Zawartość prawej ramki (o nazwie „main”) jest tworzona dynamicznie przez odwołania do danych pochodzących z serwera. Kolejnym zbiorem jest kwocinana.html. Zbiór ten jest wykorzystywany w momencie odwoływania się do tezausa.

Oczywiście ustawienia zbiorów są arbitralne i można je zmieniać na własny użytek, pamiętając o właściwych odwołaniach.

Jednak, gdy użytkownik zamierza zaimplementować system w standardowy sposób, konieczne są pewne minimalne zmiany:



1. Zamienić odwołania w zbiorze index.html dla zmiennych „database=...” (nazwa bazy) i „table=...” (nazwa tablicy wyszukiwań uproszczonych) na **<HTML>**

Należy znaleźć w zbiorze index.html następujący wiersz:

```
<FRAME SRC="/cgi-bin/isis3w.exe?database=demo&search_type=entry&lang=eng&table=mons" NAME="display">
```

i zmienić odwołania w miejscach wytłuszczonych. W przypadku zaadaptowania systemu do potrzeb tworzonej aplikacji bazy AGRIN wiersz powinien wyglądać następująco:

```
<FRAME SRC="/cgi-bin/isis3w.exe?database=agrin&search_type=entry&lang=eng&table=mons" NAME="display">
```



2. Zmienić odwołania w zbiorze toc.html dla zmiennej „database=...” (nazwa bazy) i „table=...” (odpowiednio dla tablicy wyszukiwań-uproszczonych i zaawansowanych)

<HTML>

Przy założeniu, że nasza tablica wyszukiwań prostych została zdefiniowana jako mons, a wyszukiwań złożonych jako mona, w zbiorze toc.html należy zmienić następujący wiersz:

Dla wyszukiwania uproszczonego:

```
<A HREF="/cgibin/isis3w.exe?database=demo&search_type=entry&page_header=ephmon&lang=eng&table=mons" TARGET="display" onmouseover="imgON('btn1')"
```

Dla wyszukiwania zaawansowanego:

```
<A HREF="/cgibin/isis3w.exe?database=demo&search_type=entry&page_header=ephmon&lang=eng&table=mona" TARGET="display" onmouseover="imgON('btn2')"
```



3. Implementacja tezauryasa

**CDS/ISIS
FORMATY**

Przy założeniu, że w naszej bazie są angielskie odpowiedniki terminów wyszukiwawczych, możemy użyć istniejącego tezauryasa AGROVOC. Jeśli zdecydujemy się zaimplementować funkcje wyszukiwania przez tezaurus, należy wykonać następujące czynności:

- skopiować bazę tezauryasa z CD-ROM,
- w formatach EK2.PFT, EA2.PFT oraz ER2_0.pft znaleźć ciąg: „/demo/” i zamienić go na odpowiadający nazwie katalogu tworzonej aplikacji.

Jeśli zaś zdecydowaliśmy się korzystać z innego tezauryasa (np. z makrotezauryasa OECD): należy skopiować bazę tezauryasa ustalając nazwę odpowiednią do ustawień stosownych parametrów ISIS (syspar.par, por. Aneks B). Następnie trzeba sprawdzić strukturę danych. Można również dokonać tłumaczenia tezauryasa do postaci wymaganej dla bazy THES; część angielska tezauryasa OECD powinna być kompatybilna ze strukturą wymaganą przez istniejące formaty (więcej w 4.7.3).

Zachowaj FST z aplikacji DEMO i przeprowadź pełną generację zbioru odwróconego.

W formatach EK2.PFT, EA2.PFT oraz ER2_0.pft znaleźć ciąg: „/demo/” i zamienić go na odpowiadający nazwie katalogu tworzonej aplikacji.

Wszelkie poważniejsze zmiany w zbiorach: index.html oraz kwocinana.html są zalecane jedynie w przypadku dobrej znajomości języka HTML.

4.6.2. Zbiory HTP

Pamiętajmy, że wszystkie formularze wyświetlania danych są tworzone automatycznie przez specjalne zbiory HTP¹¹. Zbiory HTP są wzorami dla dynamicznie tworzonych przez program cgi dokumentów HTML. Zdefiniowano następujące wzorce:

- **search.htp** – wzorce wykorzystywane w celu budowania formularzy wyszukiwawczych;
- **result.htp** – wzorce do tworzenia stron z wynikami wyszukiwań;
- **thes.htp** – wzorzec do tworzenia strony z polem do wprowadzania deskryptorów wyszukiwawczych z tezaursusa;
- **view.htp** – wzorzec do tworzenia strony historii wyszukiwań oraz strony do wyszukiwania w języku wyszukiwawczym ISIS;
- **browse.htp** – wzorzec do przeglądania bazy ISIS.

W przypadku standardowej instalacji systemu nie ma potrzeby modyfikacji zbiorów HTP. Jedynym plikiem jaki można zmienić, jest search.htp

W wierszu 8 zmieniamy tytuł aplikacji: ISIS3W DEMO by CC and ICIE.

W wierszu 50 zmieniamy: Full Library Catalogue.

Podobnie zmieniamy wiersze 180-103 zawierające tekst ze wskazówkami, jak korzystać z systemu.

Dodatkowo w szczególnych przypadkach modyfikacje mogą dotyczyć kolorów, własności tabeli, tekstu opisującego poszczególne elementy itd.

Możliwe jest również zastosowanie kaskadowych arkuszy stylów, które są najlepiej widoczne w przeglądarce Internet Explorer.

4.7. Udoskonalanie interfejsu

Możemy założyć, że większość funkcji systemu została omówiona. Pozostają jednak pewne kwestie, których znajomość może okazać się bardzo pożyteczna i którym poświęcamy ten podrozdział.

4.7.1. Funkcja przeglądania danych

Przeglądanie danych jest jedną z użytecznych funkcji systemu WWW-ISIS i polega na wyświetleniu rekordów z bazy danych w porządku ustalonym przez zbiór odwrócony. Proces przeglądania rozpoczyna się wraz z wyświetleniem formularza zawierającego pola. W momencie, kiedy użytkownik wybiera jedną ze

¹¹ Więcej informacji na temat zbiorów HTP znajdzie czytelnik w dokumentacji systemu WWW-ISIS.

zdefiniowanych wcześniej metod przeglądania danych (zbiór `browse.def`. [skrót_języka_eticznego]) program `isis3w.exe` wykonuje następujące czynności:

- otrzymuje prefiksy wskazujące, która część zbioru odwróconego powinna być wyświetlona oraz wybiera format ustalony w `browse.def`;
- zaczynając od wskazanego terminu (lub ciągu znaków) program wybiera kolejne terminy ze zbioru odwróconego i dla każdego terminu odczytuje punkty dostępu, rekordy wskazujące i formatujące.

Zwykle funkcja przeglądania może być wykorzystana, jeśli jesteśmy w stanie wyróżnić pola pozwalające na przeglądanie bazy w określonym porządku. W przypadku baz bibliograficznych, do tego celu można wykorzystać pola typu: autor, tytuł lub ISBN. Funkcja przeglądania danych jest bardzo pomocna zwłaszcza dla niedoświadczonych użytkowników. W aplikacji DEMO wykorzystywana jest do przeglądania bazy tezaury (porządek alfabetyczny deskryptorów, indeks KWOC lub porządek systematyczny).

Parametry kontrolujące proces przeglądania danych są zdefiniowane w tabelach definiujących przeglądanie danych WWW-ISIS, są one podobne do tabeli definiujących wyszukiwanie danych WWW-ISIS (por. pkt 4.5.2.1.). Ich parametry przechowywane są w zwykłych zbiorach tekstowych, które podzielone są na odpowiednie sekcje. Każda sekcja opisuje jeden typ rekordu. Typy te mogą pochodzić z jednej bazy danych, mogą też należeć do różnych baz ISIS.

Sekcja składa się z następujących komponentów:

- wiersza definiującego typ rekordu;
- przedrostka oraz formatu listy wyboru;
- definicji scalającej (opcjonalne).

4.7.1.1. Wiersz definiujący typ rekordu

Wiersz określający typ rekordu ma taką samą strukturę, jak w tablicach definicji wyszukiwania WWW-ISIS. Jest umieszczany w kwadratowych nawiasach i wygląda następująco:

```
[typ_rekordu; ISISowy_typ_rekordu; nazwa_na_formularzu]
```

gdzie:

`typ_rekordu` jest typem rekordu używanym przez program cgi;

`ISISowy_typ_rekordu`; jest nazwą typu rekordu;

`nazwa_na_formularzu` jest tekstem wyświetlanym użytkownikowi w formularzu przeglądania danych.

Poniżej przedstawiony jest przykład typu rekordu zdefiniowany dla bazy THES (tezaury):

```
[dsc1; DS1; THESAURUS Terms]
```

4.7.1.2. Lista prefiksów

Jak już zostało wcześniej wspomniane, proces przeglądania danych polega na przeglądaniu bazy w ustalonym porządku tak, jak jest to ustalone w zbiorze odwróconym. Użytkownik widzi pole typu kombi z różnymi opcjami (patrz rys. 52). W momencie wyboru danej opcji do programu ISISDLL przekazywana jest informacja o przedrostkach, co w następstwie powoduje wyświetlenie zbioru odwróconego i odpowiednich rekordów ze zbioru głównego. Lista opcji jest zde-

finiowana w tablicy przeglądania danych przez wiersz ustalający prefiks. Każdy wiersz ma następującą postać:

```
[prefix; ISISowy-prefiks * nazwa-formatu; wyświetlana_nazwa]
```

ISISowy-prefiks składa się z prefiksu takiego, jaki został zdefiniowany w zbiorze odwróconym (przez FST);

* – stały element (pojawia się po prefiksie zdefiniowanym w zbiorze odwróconym);

nazwa-formatu – nazwa zbioru formatującego (oczywiście poprzedzonej znakiem „@”);

wyświetlana_nazwa – jest nazwą widoczną na formularzu.

Dla bazy THES została przygotowana następująca przykładowa lista (zapisana w zbiorze browse.def.eng):

```
[prefix; ES:*@EK1; KWOC]
[prefix; E:*@EA1; alphabetic]
```

W powyższym przykładzie KWOC tworzony jest przy wykorzystaniu prefiksu ES: oraz formatu EK1.PFT, zaś alfabetyczny spis deskryptorów wykorzystuje prefiks E: oraz format EA1.PFT.

4.7.1.3. Lista wyboru formatów

Lista wyboru formatów wyświetlania danych wygląda tak samo jak analogiczna w tablicach definicji wyszukiwania WWW-ISIS (por. 4.5.2.). Konieczne jest zdefiniowanie co najmniej jednego formatu. Pod koniec pojawia się dodatkowy wiersz (który znajdzie zastosowanie w późniejszych wersjach systemu):

```
"x", "x", "x", "1"
```

Poniżej przedstawiona jest lista wyboru dla bazy THES:

```
[format; @EA1; alphabetic]
```

4.7.1.4. Definicje powiązań

Definicje powiązań (*binding definition*) są szczególnie cenne, gdy zamierzamy połączyć rekordy z dwóch różnych baz danych (pobierane dane służą jako wyrażenia wyszukiwawcze w innej bazie). Zastosowanie tego typu połączeń jest szczególnie użyteczne w przypadku tezaury, bowiem wybierając odpowiednie deskryptory możemy zastosować je jako terminy wyszukiwawcze w innej bazie danych.

Forma zapisu jest następująca:

```
"binding","nazwa_powiązania_i","baza_danych","typ_rekordu","pole"
```

Można zastosować wiele różnych definicji powiązań, co w konsekwencji oznacza, że jeden tezaurus może być zastosowany do przeszukiwania wielu baz danych.

Przykład łączenia tezaury z bazą DEMO (przechowywanego w bazie THES), dla typu rekordu **mona** oraz pola **mthesaurus** wygląda następująco:

"binding", "bind1", "DEMO", "mona", "mthesaurus"

Poniżej przedstawiona jest tablica definicji przeglądania danych WWW-ISIS. Przeglądanie danych dotyczy bazy THES. Baza tezauryusa może być przeglądana w porządku alfabetycznym deskryptorów i/lub jako tzw. KWOC.

Tabela 11
Tablica przeglądania (browse.det.eng) WWW-ISIS dla bazy THES

1.	# created by gjb 11/10/99	Komentarz
2.	# THESAURUS based search	
3.	[dsc2; DS2; THESAURUS Terms]	Definicja typu rekordu
4.	[prefix; ES:*@EK2; KWOC]	Przedrostek dla pól kombi pojedynczego wyboru
5.	[PREFIX; E:*@EA2; ALPHABETIC]	
6.	[format; @EA2; alphabetic]	Definicja wyboru formatu wyświetlanych danych
7.	"x", "x", "x", "1"	Dla późniejszych zastosowań
8.	"binding", "bind2", "DEMO", "mona", "mthesaurus"	Definicja powiązań z innymi bazami

4.7.2. Jednoczesne przeglądanie wielu baz

Funkcja ta umożliwia przeglądanie wielu baz jednocześnie w formie pojedynczej tabeli. Użytkownik ma możliwość wyszukiwania w wielu bazach jednocześnie postrzegając je jako jedną bazę. Konieczne jest spełnienie następujących warunków:

- We wszystkich przeszukiwanych bazach należy używać tych samych prefiksów dla danego klucza wyszukiwawczego.
- Dla wszystkich baz należy zastosować te same formaty.
- W zbiorze BROWSE.DEF.(+skrót języka) należy dodać odpowiednią linię do każdej tabeli (por. przykład poniżej):

[databases; db1;db2;db3]

Jeśli zamierzamy wprowadzić przeglądanie np. dla bazy katalogów i bazy akwizycji (np. o nazwach faobib oraz acquib), należy w tabeli przeglądania dodać następujący wiersz (patrz wytłuszczenie):

[monb; MON; FAO Online Catalogue]
[prefix; TI=*@ELBTI; Title]
[prefix; CO:*@ELBCO; Conference]
[prefix; ST:*@ELBST; Serial Title]
[format; @ELBTI; Short]
[databases; faobib;acquib]
"x", "x", "x", "1"

Ścieżki dostępu do baz FAOBIB oraz ACQUIB pobierane są ze zbioru dbn.def (por. Aneks A). W zbiorze tym przechowywane są informacje o ścieżkach dostępu do baz danych (patrz tab. 12). Oczywiście wszystkie formaty określone w tablicy definicji przeglądania danych będą funkcjonowały jako pojedyncze wyszukiwanie, poza tym, że dla każdej bazy danych format dla niej właściwy jest pobierany z odpowiedniego dla tej bazy katalogu. W przypadku naszych baz FAOBIB oraz ACQUIB oznacza, że ich formaty to odpowiednio: ELBTI, ELBCO, ELBST.

Przykład zbioru dbn.def Tabela 12

Nazwa bazy w konwencji CDS/ISIS	Pełna ścieżka dostępu do bazy ISIS
FAOBIB	C:\Xitami\Webpages\KnowledgeManagement\DB\most\
ACQUIB	C:\Xitami\Webpages\KnowledgeManagement\DB\client\
ELBTI	C:\Xitami\Webpages\KnowledgeManagement\DB\users\
ELBCO	C:\Xitami\Webpages\KnowledgeManagement\DB\comp\
ELBST	C:\Xitami\Webpages\KnowledgeManagement\DB\event\

4.7.3. Implementacja tezauryasa

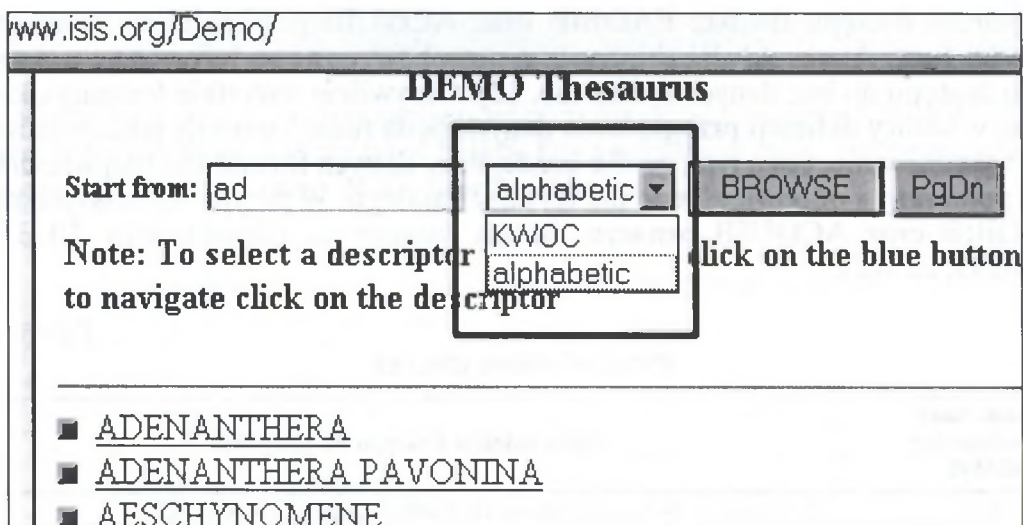
Przypomnijmy, że WWW-ISIS daje możliwość zastosowania dowolnego tezauryasa w procesie wyszukiwania. Użytkownik może przeszukiwać bazę tezauryasa (alfabetycznie, systematycznie bądź przez tzw. KWOC), a znalezione deskryptory stosować jako terminy wyszukiwawcze w bazie danych. Terminy te zbierane są w specjalnym polu tekstowym, gdzie mogą być bezpośrednio wysyłane jako zapytanie do bazy lub poddawane edycji w celu stworzenia dowolnego zapytania wyszukiwawczego.

W obecnej wersji systemu WWW-ISIS zastosowano w aplikacji DEMO tezaurus AGROVOC. Najprostsza metoda implementacji tezauryasa AGROVOC została omówiona w podrozdziale 4.6.1. W przypadku dowolnego innego tezauryasa konieczne jest uczynienie struktury bazy tezauryasa (jako bazy ISIS) zgodną ze strukturą bazy THES. Tablica wyboru pól (FST) oraz tablica selekcji pól (FST) są dostępne jako aneks (patrz Aneks F).

Pamiętajmy, że tworząc lub zmieniając strukturę bazy powinniśmy się kierować następującymi regułami:

- 1. Jeden rekord powinien zawierać wszystkie niezbędne informacje dot. jego powiązań z innymi terminami (terminy kojarzeniowe, szersze, węższe itd), działów itd.
- 2. Dla każdego typu relacji powinien być zastosowany osobny niepowtarzalny identyfikator pola.
- 3. Każda relacja powinna być określona przez pole powtarzalne.
- 4. Wszystkie pola odwołujące się do innych terminów przez daną relację powinny być ułożone w porządku alfabetycznym.

Podstawową kwestią pozostaje: jak poruszać się w bazie tezauryasa (rys. 52).



Rys. 52. Nawigacja w obrębie tezauryza

Zwykle możemy zaimplementować następujące rodzaje nawigacji przez:

- Indeks alfabetyczny.
- Indeks KWOC.

Oba sposoby mogą być zastosowane przez wykorzystanie tzw. funkcji przeglądania danych. Dotychczas udało nam się zastosować opcje 1 oraz 2.

Jednakże może okazać się, że zapagniemy oprócz bazy deskryptorów przeglądać bazę askryptorów (w FDT bazy THES pole o znaczniku 8, por. Aneks F).

Tak więc główny wiersz tablicy definicji pól (FST) dla indeksu alfabetycznego powinien wyglądać następująco:

```
1 0 "E:"V1
8 0 "E:"V8
```

Zaś w przypadku indeksu KWOC FST (tablica selekcji pól) powinna zaindeksować każde słowo osobno, a więc stosujemy technikę selekcji 8:

```
1 8 '/ES:/',V1
8 8 '/ES:/',V8
```

Tablice definiujące przeglądanie danych mogą być również wykorzystywane do przeglądania tezauryza. Zawierają one sekcję łączącą bazę tezauryza z daną bazą ISIS. Przypomnijmy, że mechanizm łączenia umożliwia właściwe funkcjonowanie tzw. schowka (miejsca w pamięci), w którym przechowywane są wybrane terminy. Aby wprowadzić znalezione w tezaurysie terminy do standardowego formularza wyszukiwawczego, program cgi musi „znać” nazwę bazy, tablicy oraz pola. Powtarzanie tego typu danych w każdym połączeniu typu „add_term” jest nadmiarowe, dlatego stosuje się tzw. binding_name. Określona informacja dotycząca łączenia baz (binding_name) zawarta jest w tablicy definiującej przeglądanie danych.

4.7.4. Więcej o formatowaniu w WWW-ISIS

W niniejszym podrozdziale zawarte są dodatkowe cechy języka formatowania charakterystyczne dla pakietu WWW-ISIS.

4.7.4.1. Format „exits”

Ponieważ nie ma możliwości zastosowania typowych dla języka Pascal-ISIS funkcji exit, konieczne staje się wprowadzenie specjalnych wirtualnych pól w celu dokonywania dodatkowych operacji na danych. Poniżej przedstawiona jest lista tzw. wirtualnych pól, które mogą dostarczać dodatkowych informacji niedostępnych *explicite* w bazie.

9980 – termin nadrzędny w grupie przeglądanych terminów;

9982 – położenie prefiksu w dokumencie;

9984 – nazwy baz wykorzystywane w wyszukiwaniu;

9991 – wymagana liczba wyników wyszukiwań;

9999 – wyświetlona liczba wyników wyszukiwań.

Powyższe znaczniki mogą być używane w języku formatowania w celu przeglądania danych, definiowaniu nagłówków itd.

4.7.4.2. Funkcja ISISREF

Istniejąca wersja ISIS-DLL nie daje możliwości poprawnego korzystania z funkcji REF i dlatego konieczne jest zastosowanie pewnych rozwiązań rekurencyjnych. Polegają one na tworzeniu w formacie łączy ISISREF. W tym celu format w WWW-ISIS winien mieć następującą postać:

```
<ISISREF DBN=nazwabazydanych; MFN=nnn; PFT=@nazwaformatu>  
<ISISREF DBN=nazwabazydanych; L=punkt_dostępu; PFT=@nazwaformatu>
```

W pierwszym przypadku połączenie wskazuje na inny rekord (przez MFN numer rekordu), zaś w drugim przypadku połączenie wskazuje jeden lub więcej rekordów dzięki pętli funkcji L wskazującej na zbiór odwrócony skąd pobierane są odwołania do rekordów.

```
<ISISREF DBN= CDS; MFN=120; PFT=@CDS>
```

W tym przypadku program WWW-ISIS będzie próbował otworzyć rekord o numerze 120 i zastosować do niego format CDS

Przykład 7

```
<ISISREF DBN= FAOLEX; L=PD=LEX-000011; PFT=@title>
```

Program będzie odwoływał się do punktu dostępu PD=LEX-000011 i będzie próbował sformatować go zgodnie z formatem title.



Analizując bardziej szczegółowo powyższą funkcję zauważymy, że podane połączenie:

- dokonuje ekstrakcji połączenia z formatowanego obszaru danych;
- wybiera parametry połączeń i wedle nich tworzy odpowiedni tekst;
- jeśli ten tekst nie zawiera już żadnego połączenia, wstawiany jest on w miejscu połączenia. Jeśli znajdowane jest połączenie, operacja jest powtarzana.

Jeśli funkcja pętli wskazuje na połączenie, to może okazać się, że istnieje więcej niż jedno połączenie. W takim przypadku stosowany jest format wskazany przez parametr PFT i kolejne połączenia są zamieniane przez odpowiedni tekst.

4.7.5. Przechowywanie zakładek w celu powtórzenia wyszukiwania

System daje też możliwość zapamiętywania zapytań. Wywoływane jest to przez przycisk, który powinien być zdefiniowany w zbiorze http – search.http. Zatem wiersz 50 winien dodatkowo zawierać:

```
<INPUT TYPE=BUTTON NAME="ADD_BOOKMARK" VALUE="SAVE" onClick="window.open(makeurl('/cgi-bin/isis3w.exe?search_type=view_query_current'),'View_Query','width=610,height=500,location=no,status=no,toolbar=yes,menubar=yes,scrollbars=yes,resizable=yes')">
```

Powyższa funkcja pozwala na zachowanie danego zapytania w formie zakładki:

1. Kiedy użytkownik zdecyduje się na zachowanie zapytania, przyciska przycisk **SAVE QUERY**.
2. System wyświetla okno, w którym jest wyświetlone zapytanie w formie języka zapytań CDS/ISIS. Okno to może być wykorzystane jako zakładka.
3. Za każdym razem, jeśli będziemy chcieli powrócić do zapytania, wystarczy wybrać je z listy zakładek.

4.7.6. Funkcja eksportu danych

Funkcje eksportu danych są dostępne na poziomie języka formatowania (możliwe jest stworzenie formularza pozwalającego na dokonanie eksportu wybranych rekordów). Można dokonać eksportu znalezionych rekordów jako zbioru ISO2709 lub jako zbioru tekstowego, RTF lub XML.

Możliwe jest zaimplementowanie w formularzu wyświetlania danych funkcji pozwalającej na eksport danych. Aby zaimplementować tę funkcję w formularzu wyszukiwawczym, należy posiadać odpowiednią wersję zbioru (szablону) http oraz dokonać odpowiednich zmian w formatach wyświetlania danych.



1. RESULT.HTP (PO VI 2001)

Należy wykorzystać szablon result.http zgodny z wersją isis3w.exe (utworzony po czerwcu 2001).



2. ZMIEN FORMATY WYŚWIETLANIA

CDS/ISIS FORMATY

W odpowiednich formatach wyświetlania danych wprowadzić następujący wiersz HTML (pola typu zaznacz= CHECKBOX):

```
'<TABLE WIDTH=100%><TR><TD WIDTH=4% VALIGN=TOP><INPUT  
TYPE="CHECKBOX" NAME="selmfn" VALUE="dbn: ',mfn, '"></TD>'>#
```

Jest możliwe zdefiniowanie odpowiednich ekranów Eksport/Import w szablonie result.http (wiersz 89-96).

OGRANICZENIA:

Obecna wersja ISIS-DLL nie daje możliwości wykorzystania zbioru FST jako tablicy konwersji przy importowaniu danych z formatu ISO-2709. Inna funkcja eksportu danych dostępna jest z menu. W przypadku, kiedy zamierzamy dokonać eksportu/importu danych, możemy wybrać jedynie podzbiór rekordów. Dostępne standardy to: ISO2709, zwykły zbiór tekstowy, zbiór XML.



4.7.7. Dodatkowe funkcje wyszukiwawcze

4.7.7.1. Wyszukiwanie pełnotekstowe

System daje również możliwość zaimplementowania dodatkowo pól wyszukiwawczych pozwalających na:

- Wyszukiwanie z podaniem zakresu wyszukiwanych danych.
- Wyszukiwanie tekstowe.

Wyszukiwanie z podaniem zakresu pozwala na określenie zakresu rekordów, w których określone pola przyjmują daną wartość.

Wyszukiwanie tekstowe umożliwia zdefiniowanie zapytania w języku zapytań ISIS, zgodnie z zasadami wyszukiwania pełnotekstowego w języku zapytań ISIS.

Dla obu typów zapytań można zdefiniować osobny formularz zapytań. Typ formularza powinien zostać zdefiniowany jako 9 (type=9). Ogólna forma definicji ma następujący wygląd:

```
"Nazwa_pola_na_ekranie", "wewnętrzna_nazwa", [selector_pola], "9",  
{ "int", "text" }, { [from-par], [To-par], [Equals-par] }, length-par;
```

ZAPAMIĘTAJ:



1. Dla danego zapytania w bazie database.def można określić tylko jeden typ pola – 9 (type=9). Każda następna definicja będzie zignorowana.
2. Parametr **[selektor_pola]** jest obowiązkowy do wyszukiwania zakresowego oraz wyszukiwania typu „EQUALS”. Ma on postać wyrażenia w języku formatowania, dostarcza informacji nt. uzyskania liczby z pola, np. v83.4. Dla „wyszukiwania tekstowego” jest ignorowany.
3. Parametr „**int**” jest używany do wyszukiwania typu „range” oraz „equals”, zaś parametr „**text**” w wyszukiwaniu tekstowym.
4. **From-par** (opcjonalny) określa wyświetlany przed polem tekst podający początkowe parametry wyszukiwania.
5. **To-par** (opcjonalny) określa wyświetlany przed polem tekst podający końcowe wyszukiwanie.
6. **Equal-par** (opcjonalny) określa wyświetlany przed polem tekst podający wartości wyszukiwawcze.
7. **Length-par** ustala długość pola, do wprowadzania wartości.
8. Jeśli podany jest parametr „**text**” zamiast parametru „**int**”, należy użyć wyłącznie ostatniego parametru („**length-par**”).



Przykład 8



Założmy, że w naszym formularzu wyszukiwawczym chcemy umieścić pola pozwalające na wyszukiwanie zakresowe.

W tablicy wyszukiwań (w części poświęconej definicji wyboru pól, por. 4.5.2.3) wprowadźmy następujący wiersz:

"Year of publication: ", "yearp", "v81.4", "9", "int", "From", "To", "", "4";

Dzięki temu uzyskujemy element formularza pozwalający na wyszukiwanie zakresowe (rys. 53).

Publication year

From:

To:

Rys. 53. Formularz z polem ustalającym zakres wyszukiwań

Pola „From” oraz „To” (widoczne na rys. 53) mają długość 4 znaków (length = 4). Wyszukiwanie przez podanie zakresu wygląda następująco:

1. Standardowa część formularza wyszukiwawczego jest używana przez wyszukiwanie wykorzystujące zbiór odwrócony.

2. Następnie rezultat jest filtrowany zgodnie z warunkami ustalonymi w polach „From” i „To” i posortowany zgodnie ze specyfikacją sortowania. W podanym przykładzie rezultat porównany jest z wartością V81.4.

3. Posortowany wynik jest umieszczony na kolejnych stronach (zgodnie z ustalonymi parametrami dotyczącymi umieszczania wyników na stronach).



UWAGA:

Jeśli formularz dla standardowego wyszukiwania (wykorzystującego zbiór odwrócony) jest pusty, ignorowane są specyfikacje dotyczące zakresu.

W podobny w sposób można zdefiniować wyszukiwanie przez podanie dokładnych wartości (oczywiście w takich przypadkach parametr „from” oraz „to” powinien być pusty, podczas gdy parametr „equals” powinien zawierać odpowiedni tekst, np. „Jest równe” albo „Wynosi” itd.).



Przykład 9



Przyjrzyjmy się teraz wyszukiwaniu tekstowemu. Określenia dotyczące tekstu powinny wyglądać następująco:

"Free text search: ", "text", "", "9", "text", "", "", "", "20";

Podana wyżej definicja umożliwia wyświetlenie na formularzu pola tekstowego do wprowadzania terminów wyszukiwawczych (rys. 54).

Free text search:

Rys. 54. Wyszukiwanie pełnotekstowe

Pole tekstowe (rys. 54) składa się z 20 znaków. W polu tym można zdefiniować następujące zapytanie:

V31: 'protection' and V31: 'food'

UWAGA:

Składnia wyszukiwania tekstowego winna być spójna ze składnią wyszukiwania pełnotekstowego języka ISIS. W przypadku definiowania pola do wyszukiwania określającego zakres, parametry określające wyszukiwanie tekstowe (jeśli istnieją) są ignorowane.



4.7.7.2. Grupy And-Or

Żałujemy, że użytkownik zamierza skierować do bazy zapytanie zawierające następujące słowa „food” oraz „protection”, które winny występować w tytule, a ponadto chce mieć możliwość skorzystania z wyszukiwania pełnotekstowego. Na pierwszy rzut oka należałoby stworzyć formularz pokazany poniżej, łączący operatorem OR pole „Title” (tytuł) z polem „Serial Title” (tytuł czasopisma).

Collection:	FAO documents	AND	OR	NOT
Full text	Available	AND	OR	NOT
Search in whole record (words)		AND	Show List	
Title (words)	food;protection	AND	Show List	
Serial Title (words)	food;protection	OR	Show List	

Rys. 55. Grupy And-Or

Niestety taki formularz i w konsekwencji zapytania dadzą błędne odpowiedzi.

TI (Title *tytuł*)

ST (Serial Title – *tytuł czasopisma*)

Czyli nasz formularz spowoduje wygenerowanie następującego pytania :

FAO=FAO documents * S0=Available
("TI=food"*"TI=protection")+ ("ST=food"+"ST=protection")



Zaś naszą intencją było skierowania następującego zapytania do bazy:

FAO=FAO documents * S0=Available *
(("TI=food"*"TI=protection")+("ST=food"+"ST=protection"))



Aby uniknąć lub co najmniej zminimalizować możliwość popełniania takich błędów, użytkownik winien zdefiniować grupy AND/OR, które są grupami pól ujętymi przez system w nawiasy.

W przypadku systemu wyszukiwawczego FAOBIB można dokonać wyboru pewnych pól, które pełnią rolę filtrów jak np. *język, wyszukiwanie tekstowe, zbiór* i zdefiniować je jako jedną grupę (ujętą w nawiasy), zaś pozostałe pola jako kolejną (inną) grupę AND/OR. Po zdefiniowaniu obu osobnych grup oddzielnie możemy uniknąć ewentualnych błędów w wyszukiwaniu.

Dla każdego formularza wyszukiwawczego można zdefiniować do dwóch grup. Grupa może być zbudowana z towarzyszących pól. Aby zdefiniować grupę należy zastosować następującą składnię:

```
[group; nazwa_pola_początkowego; nazwa_pola_końcowego]
```

gdzie:

nazwa_pola_początkowego oraz *nazwa_pola_końcowego* są używanymi przez system nazwami pól tak jak zostało one zdefiniowane w podanym formularzu

4.8. Ćwiczenia



Część z przedstawionych poniżej ćwiczeń wymaga zrozumienia środowiska sieci oraz zaznajomienia się obsługą serwerów WWW, co jest w miarę szczegółowo omówione w Aneksie E.



Ćwiczenie 14. Tworzenie FST dla bazy CDS

CDS/ISIS
FST

Utworzyć właściwy dla systemu WWW-ISIS zbiór FST dla bazy CDS dostarczonej z pakietem CDS/ISIS. Poniżej znajduje się uproszczona FDT dla bazy CDS

Conference main entry	npdz	12 300 0 0
Title	z	24 500 0 0
Edition		25 100 0 0
Imprint	abc	26 300 0 0
Collation	abc	30 100 0 0
Series	vz	44 300 0 1
Notes		50 500 0 0
Keywords		69 1000 0 0
Personal Authors		70 100 0 1
Corporate Bodies		71 300 0 1
Meetings	npdz	72 300 0 1
Added Title	z	74 500 0 1
Other language titles	z	76 500 0 1

Ćwiczenie 15

Dla bazy CDS zaprojektuj i przetestuj tablice wyszukiwań w WWW-ISIS zgodną z podanym FST:

CDS/ISIS
FST



```
1 0 'T=BIB'  
70 0 MHU, (|A=|V70/)  
24 8 '/TI=/',V24  
69 6 '/D=/'V69  
26 0 "Y=",V26^c  
71 8 '/C=/' V71+|; |  
44 8 '/S=/' V44+|; |
```

Ćwiczenie 16. Formaty wyświetlania

Dla bazy CDS sporządź dwa formaty wyświetlania danych: skrócony oraz pełny. Skrócony powinien zawierać połączenie pozwalające na przejście do pełnego rekordu. Sporządź pełny format w formie tabeli.

WWW-ISIS



Ćwiczenie 17. Formaty wyświetlania + testy

Przetestuj na komputerze ćwiczenie 16.



Ćwiczenie 18

Umieść i przetestuj w formularzu wyszukiwawczym pola pokazane na rys. 53. Zastosuj następujący zapis:

"Year of publication: ", "yearp", "v81.4", "9", "int", "From", "To",
"", "4";

Ćwiczenie 19. Zmiana FST

Do bazy DEMO dodaj opcję przeglądania według autorów oraz tytułów (należy zmienić FST oraz zbiór browse.def.eng).

CDS/ISIS
FST

WWW-ISIS



Ćwiczenie 20. Wyszukiwanie na CD-R (por. Aneks E)

Jeśli posiadasz nagrywarkę CD-R zainstaluj serwer Sprite wraz ze stworzoną przez siebie aplikacją WWW-ISIS (zadanie do wykonania po zakończeniu rozdz. 5).



Ćwiczenie 21. Tłumaczenie interfejsu

Postaraj się przetłumaczyć cały interfejs wyszukiwawczy bazy DEMO na język polski zachowując jego funkcjonalność.

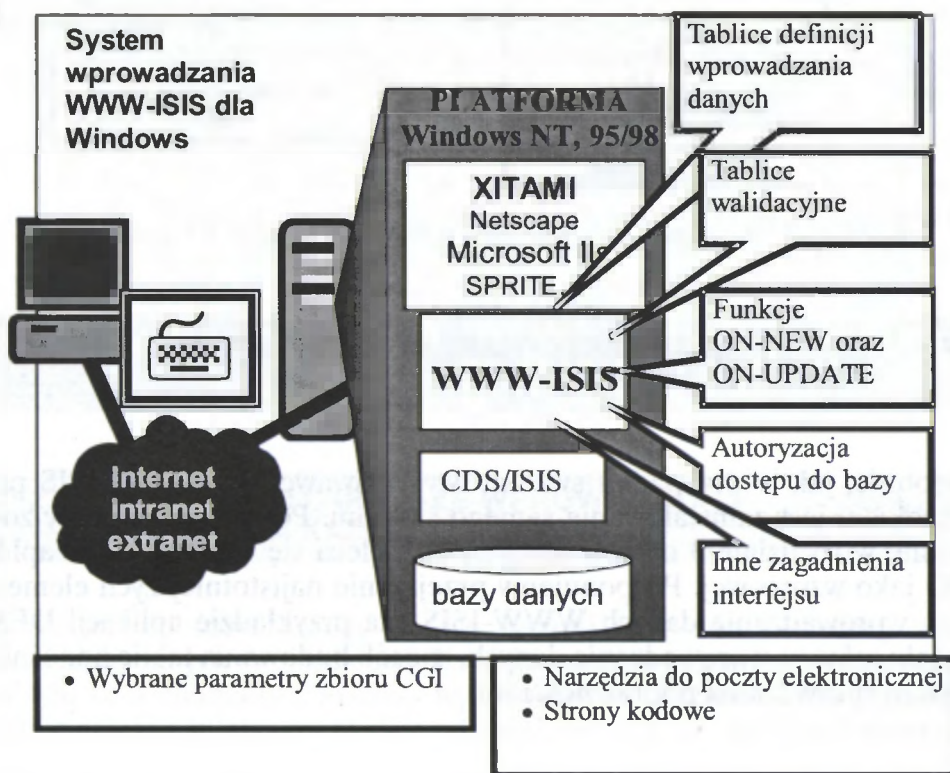
WWW-ISIS



5. Wprowadzanie danych w WWW-ISIS

Część ta poświęcona jest tym, którzy zamierzają udostępniać bazy danych ISIS w sieci Internet/intranet w celu dokonywania zmian w bazie (wprowadzania nowych rekordów, uaktualniania, kopiowania oraz usuwania). Chociaż rozdział ten stanowi w dużym stopniu część samodzielny i nie wymaga zaznajamiania się z częścią 4 dotyczącą wyszukiwania danych, to jednak pewne elementy z rozdziału 4 mogą okazać się pomocne w konstruowaniu w pełni funkcjonalnego interfejsu wprowadzania danych. Zakładamy, że po zapoznaniu się z przedstawioną w niniejszym rozdziale całością materiału, czytelnik będzie w stanie stworzyć swój własny interfejs wprowadzania danych.

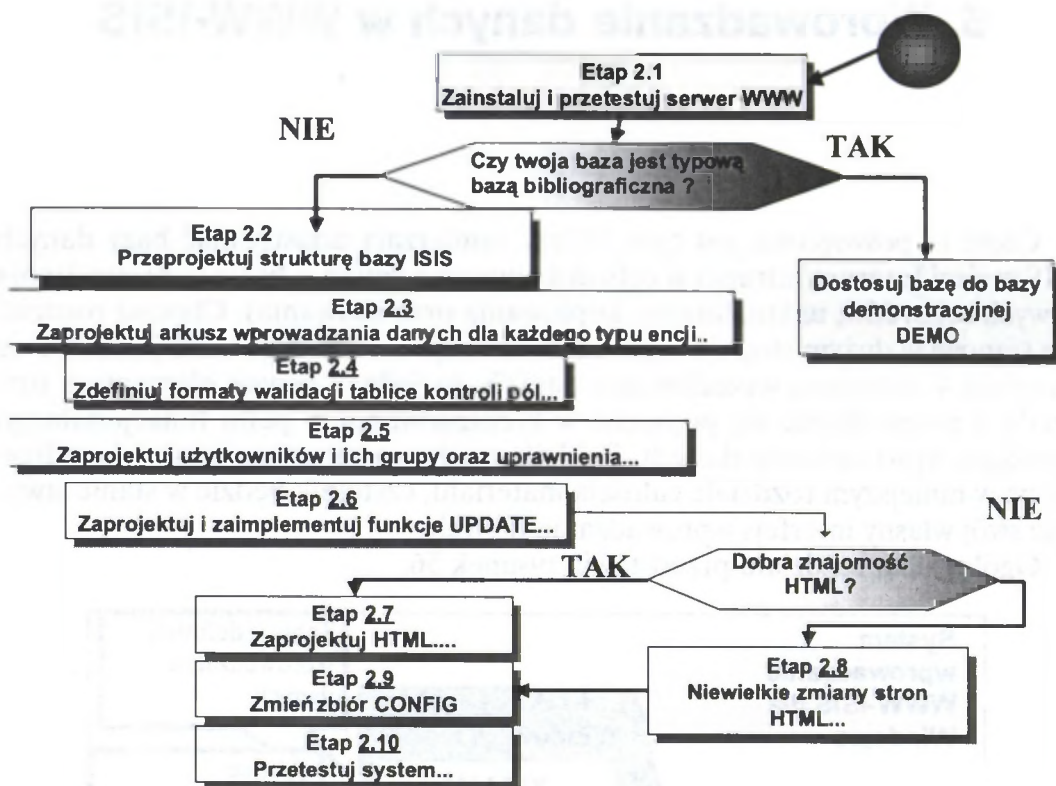
Ogólny plan rozdziału przedstawia rysunek 56.



Rys. 56. Omówienie systemu wprowadzania danych WWW-ISIS

5.1. Plan działania

Po zainstalowaniu serwera WWW oraz systemu WWW-ISIS sugerowane są czynności przedstawione na rys. 57. Omówmy po kolei poszczególne z nich, zaczynając od etapu 2.1. a kończąc na 2.8.



Rys. 57. Schemat implementacji systemu WWW-ISIS (wprowadzanie danych)



ETAP 2.1. INSTALACJA I TESTOWANIE SERWERA WWW I APLIKACJI DEMO

Podobnie, jak w przypadku systemu wyszukiwawczego WWW-ISIS pierwszym krokiem jest zainstalowanie samego systemu. Poszczególne etapy zostały omówione w rozdziale 3 oraz w Aneksie E. Zaleca się wykorzystanie aplikacji DEMO jako wzorcowej. Proponujemy przejrzanie najistotniejszych elementów systemu wprowadzania danych WWW-ISIS (na przykładzie aplikacji DEMO) takich jak: arkusze wprowadzania danych, sposób budowania tablic odnośników, podsystem sprawdzania poprawności itd.



REJESTRACJA W SYSTEMIE

Nazwa użytkownika:

User-ID =

dba

Hasło użytkownika:

Password =

dba

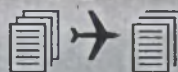
REGUŁY WPROWADZANIA DANYCH

Nie używaj nawiasów trójkątnych < >
Dla pól powtarzalnych można użyć średnika ;



WSKAZÓWKA:

Jeśli nasza baza jest **typową bazą bibliograficzną**, proponuje się zaadaptowanie istniejącej wersji DEMO systemu. W przeciwnym przypadku postępujemy zgodnie ze schematem na rys. 57.



ETAP 2.2. PRZEPROJEKTOWANIE I/LUB ZAPROJEKTOWANIE STRUKTURY BAZY ISIS

CDS/ISIS
FST FDT



Generalnie, przeprojektowanie bazy powinno uwzględniać jej funkcjonalność w systemie WWW-ISIS (czyli w naszym przypadku możliwość wprowadzania danych). Należy też rozważyć, które encje powinny być umieszczone w bazie ISIS i jak powinny wyglądać relacje między nimi. Następnie należy wyodrębnić pola, według których będziemy dokonywali przeszukiwania bazy, a w konsekwencji przeprojektować tablicę FST uwzględniając planowane punkty dostępu oraz utworzenie indeksów dla wszystkich przeszukiwanych pól wyróżnionych przez prefiksy.

ETAP 2.3. ZAPROJEKTUJ ARKUSZE WPROWADZANIA DANYCH DLA KAŻDEGO TYPU ENCJI

WWW-ISIS



Po dokonaniu rozróżnienia między typami rekordów tworzymy arkusze wprowadzania danych (worksheets) dla każdego typu rekordu. Można wykorzystać w tym celu elementy z arkuszy wprowadzania danych z aplikacji DEMO. Następnie należy zaplanować tablice odnośników, arkusze dla haseł wzorcowych itd.

ETAP 2.4. ZDEFINIUJ FORMATY SPRAWDZANIA POPRAWNOŚCI

CDS/ISIS
FORMATY



W celu osiągnięcia funkcjonalności systemu, należy zdefiniować formaty sprawdzania poprawności wprowadzanych rekordów (do sprawdzania poprawności na poziomie rekordu – poprawności globalnej) oraz tablice kontroli pól (do sprawdzania poprawności na poziomie pola, tj. sprawdzania zawartości pola), wprowadzić ustawienia dotyczące wartości domyślnych pól oraz pozostałe formaty.



ETAP 2.5. ZAPROJEKTUJ UŻYTKOWNIKÓW I ICH GRUPY ORAZ UPRAWNIENIA

WWW-ISIS

W celu właściwego funkcjonowania systemu konieczne jest zaprojektowanie grup użytkowników i uprawnień dostępu do danych.



ETAP 2.6. ZAPROJEKTUJ I ZAIMPLEMENTUJ FUNKCJĘ AKTUALIZACJI REKORDÓW (UPDATE)

**CDS/ISIS
FORMATY**

Na tym etapie projektuje się i implementuje funkcje aktualizacji rekordów w bazie danych, metody dostępu do rekordów, które chcemy aktualizować (przeglądanie i wyszukiwanie), formaty wyświetlania, przyciski funkcyjne itd.



ETAP 2.7. ZAPROJEKTUJ HTML

<HTML>

Jeśli twoja znajomość języka HTML nie jest najlepsza, nie zaleca się szczególnego przeprojektowywania istniejących struktur HTML (dotyczy zbiorów: index.html, toc.html, de.html itd.).



ETAP 2.9. ZMIEN ZBIÓR KONFIGURACYJNY (ISIS3W.EXE.CFG)

WWW-ISIS

Czasami okazuje się konieczne dokonanie niewielkich zmian w zbiorze konfiguracyjnym WWW-ISIS (isis3w.exe.cfg patrz Aneks A).



ETAP 2.10. TESTUJ SYSTEM

Jak zwykle najistotniejszym etapem prac jest dokładne przetestowanie systemu, nawet jeśli wydaje się nam, że wszystkie etapy jego tworzenia przebiegły pomyślnie.

Zanim w następnych podrozdziałach omówimy szczegółowo etapy implementacji systemu wprowadzania danych WWW-ISIS, przedstawmy pokrótce wstępną fazę jego projektowania.

5.2. Wstępna faza projektowania bazy

W procesie budowania systemu wprowadzania danych WWW-ISIS należy odróżnić fazę projektowania od fazy implementacji systemu. Zanim przejdziemy do konkretnych etapów implementacji systemu, poniżej przyjrzymy się nieco dokładniej metodyce projektowania systemu wprowadzania danych WWW-ISIS.

KROK 1 – ZDEFINIOWANIE TYPÓW REKORDÓW

W momencie projektowania bazy ISIS, ważne jest zdefiniowanie typów rekordów oraz relacji między nimi.

KROK 2 – OKREŚLANIE WŁAŚCIWOŚCI PÓL

CDS/ISIS
FST FDT

Dla każdego typu rekordu należy określić właściwy podzbiór pól. Zwykle użytkownicy ISIS tworzą tablicę definicji pól (FDT) w nadziei, że jej zaprojektowanie jest wystarczające do stworzenia pełnowartościowej bazy danych. Jednak sama tablica FDT nie daje jasnych kryteriów odnośnie pól i relacji między nimi. Dlatego też konieczne jest stworzenie tabeli (tab. 13), w której należy określić właściwości poszczególnych pól.

Tabela 13

Krok 2 w projektowaniu systemu wprowadzania danych WWW-ISIS

Strona	Znacznik ISIS	Opis	Tabela odnośników		Podpola	Wymagany T/N	Powtarzalność	Reprezentacja w aplikacji WWW- ISIS
			prefiks	baza danych				
...	...	...	...	...	...	...	...	...

Dla każdego typu rekordu należy sporządzić podobną tabelę (por. tab. 13), zaś w polu „Reprezentacja w aplikacji WWW-ISIS” należy określić, w jakim arkuszu będzie się znajdowało dane pole i jaką będzie miało postać.

Zobaczmy na konkretnym przykładzie, jak będzie wyglądał niniejszy proces.



5.2.1. Przykład wstępnej fazy projektowania systemu



Zakładamy, że czytelnik potrafi ustalić odpowiednie relacje i encje, dlatego pomijamy krok pierwszy. Ponadto zakładamy, że nasza baza danych jest typową bazą bibliograficzną¹:

TRN	AA9999999	1 9 3 0
Prim. subject code	A99	7 3 3 0
Sec. subj. code		8 7 0 1
type of document		10 1 1 0
bibliogr. levels		11 3 1 0
literary indicator		12 10 1 1
Data entry status		91 1 0 0
***** a – level *****		1001 1 0 0
personal name(s) (^n^a)		100 400 0 0
corporate name(s)		110 400 0 0
English main title		120 400 0 0
French main title		122 400 0 0
Spanish main title		124 400 0 0
Other lang. main title		126 400 0 0
conference name		130 200 0 0
conference place		131 80 0 0
conference date		133 80 0 0
report/patent number		150 80 0 0
ISBN/IPC		152 80 0 0
collation		170 40 0 0
language of text		180 40 0 1
summary		181 80 0 1
notes		182 160 0 0
***** M – level *****		1002 1 0 0
personal name(s) (^n^a)		200 400 0 0
corporate name(s)		210 400 0 0
English main title		220 400 0 0
French main title		222 400 0 0
Spanish main title		224 400 0 0
Other lang. main title		226 400 0 0
conference name		230 200 0 0
conference place		231 80 0 0
conference date		233 80 0 0
edition		240 80 0 0
report/patent number		250 80 0 0
ISBN/IPC		252 80 0 0
imprint, place of publication		261 80 0 0
imprint, publisher		262 80 0 0
imprint, date of publication		263 80 0 0

¹ Struktura bazy (DEMO) jest wzorowana na strukturze bazy AGRIS.

collation	270	40	0	0
language of text	280	40	0	1
summary	281	80	0	1
notes	282	160	0	0
availability	283	160	0	0
online availability	290	160	0	0
***** S – level *****	1003	1	0	0
serial, main title	420	400	0	0
serial, sec. title	421	400	0	0
serial, ISSN	450	80	0	0
serial, date of publication	463	80	0	0
serial, collation	470	40	0	0
serial, notes	482	160	0	0
availability	483	160	0	0
online availability	490	160	0	0
***** X – level *****	1004	1	0	0
Language of indexing	500	2	0	0
En – AGROVOC	501	600	0	0
En – comments/proposed AGROVOC	502	250	0	0
Fr – AGROVOC	521	600	0	0
Fr – comments/proposed AGROVOC	522	250	0	0
Sp – AGROVOC	541	600	0	0
Sp – comments/proposed AGROVOC	542	250	0	0
lang. of local terms	700	2	0	0
local terms	701	400	0	0
lang. of abstract	710	2	0	0
abstract	714	400	0	0

Tak więc rozróżniliśmy trzy typy rekordów, patrz znaczniki: 100-182; 200-290; 420-490. W dalszej części ograniczymy się do pokazania tylko części jednego typu rekordu (artykułu z czasopisma – *paper from journal*).

Tabela 14

Parametry dla pól (arkusz wprowadzania danych)

Strona	Znacznik ISIS	Opis na formularzu (j.ang)	Prefiks w bazach odnośników	Baza danych dla tabeli odnośników (lookup tables)	Podpola	Pole wymagane T/N	Powtarzalność T/N	Reprezentacja w aplikacji WWW-ISIS
1	2	3	4	5	6	7	8	9
1	1	Temp. Record Number (TRN)	-		-	T	N	Jednowierszowe pole tekstowe
1	91	Data entry status						Lista wykluczających się opcji (przycisk radio)
1	7	Primary subj. cat. Code	CC:	valid	-	T	T	Jednowierszowe pole tekstowe

1	2	3	4	5	6	7	8	9
1	8	Secondary subj. cat. Code	CC:	valid	-	T	T	Jednowierszowe pole tekstowe
1	10	Type of publication			-	T	N	Pole typu kombi
1	11	Bibliographic level				T	N	Lista wykluczających się opcji (przycisk RADIO)
1	12	Literary indicator				N	T	Pole typu zaznacz (CHECKBOX)
2	100	Personal Author (repeatable group)			ⁿ name	N	T	Jednowierszowe pole tekstowe
					^a affiliation	N		Jednowierszowe pole tekstowe
2	110	Corporate Author	EA:	auth		N ²	T	Jednowierszowe pole tekstowe
2	120	English title	TM:	demo		N ³	N	Wielowierszowe pole tekstowe
2	122	French title	TM:	demo		N	N	Wielowierszowe pole tekstowe
2	124	Spanish title	TM:	demo		N	N	Wielowierszowe pole tekstowe
2	126	Other language title	TM:	demo		N	N	Wielowierszowe pole tekstowe
3	130	Conference name ⁴	EO:	auth		N	N	Jednowierszowe pole tekstowe
3	131	Conference place				N	N	Jednowierszowe pole tekstowe
3	133	Conference date				N	N	Jednowierszowe pole tekstowe
3	150	Report/patent number				N	N	Jednowierszowe pole tekstowe
3	152	ISBN/IPC				N	N	Jednowierszowe pole tekstowe
3	170	Collation				T	N	Jednowierszowe pole tekstowe
3	180	Language of text	LNE:	valid		T	T	Jednowierszowe pole tekstowe
3	181	Summary statement				N	N	Wielowierszowe pole tekstowe
3	182	Notes				N	T	Wielowierszowe pole tekstowe

² W rekordzie winna znajdować się informacja zarówno nt. autora osobowego, jak i zbiorowego.

³ Co najmniej jedno z pól: 120, 122, 124, 126 jest obowiązkowe.

⁴ Jeśli nazwa konferencji jest znana, to winna pojawić się również informacja o miejscu konferencji i jej dacie.

5.3. Implementacja systemu

W procesie implementacji systemu można rozróżnić 4 kroki:

KROK 1 – DEFINIOWANIE ARKUSZY WPROWADZANIA DANYCH (ENTRY.DEF)

WWW-ISIS



Wykorzystując rezultaty fazy projektowania systemu definiujemy arkusze w tablicach definiujących wprowadzanie danych. Jeśli istnieje taka potrzeba należy przeprojektować istniejące FST tak, aby było zgodne z tabelą odnośników (look-up tables) dla poszczególnych pól w arkuszach wprowadzania danych.

KROK 2 – DEFINIOWANIE TABLIC PRZEGLĄDANIA DANYCH DLA TABEL ODNOŚNIKÓW (LOOKUP TABLES)

**CDS/ISIS
FORMATY**



Definiujemy tablice przeglądania danych dla odnośników, a następnie definiujemy i testujemy formaty selekcji danych oraz testujemy arkusze wprowadzania danych.

KROK 3 – DEFINIOWANIE FORMATÓW SPRAWDZANIA POPRAWNOŚCI ORAZ AUTOMATYCZNEGO WPROWADZANIA DANYCH

**CDS/ISIS
FORMATY**



Dla każdego typu rekordów definiujemy właściwe formaty i tablice sprawdzania poprawności wprowadzanych danych⁵. Definiujemy też formaty dla automatycznego wprowadzania danych (On New oraz On Update), a następnie testujemy arkusze wprowadzania danych wraz z zaimplementowanymi funkcjami (On New oraz On Update).

KROK 4 – DEFINIOWANIE: TABLIC PRZEGLĄDANIA DANYCH (UAKTUALNIENIE REKORDÓW), FORMATÓW SELEKCJI, TABLIC WYSZUKIWAŃ (OPCJONALNIE)

**CDS/ISIS
FORMATY**



Dla każdego typu rekordów wprowadzamy procedury selekcji rekordów w celu ich aktualizacji: definiujemy tablice przeglądania danych w celu wyboru rekordów do uaktualnienia, definiujemy i testujemy formaty selekcji i opcjonalnie definiujemy tablice wyszukiwań.

⁵ Formaty sprawdzają poprawność na poziomie rekordu, tablice sprawdzają zawartość pola względem określonego zbioru wartości (tezaurusu, baz odnośnikowych).

5.4. Definiowanie arkuszy wprowadzania danych

Wszystkie arkusze wprowadzania danych są zdefiniowane w zbiorze entry.def (zwykle zlokalizowanym w katalogu głównym danej aplikacji wraz z pozostałymi zbiorami HTML)⁶. Zanim rozpoczniemy szczegółowe omawianie wszystkich parametrów tego zbioru, opiszmy w skrócie ogólną zasadę definiowania arkuszy wprowadzania danych i przyjrzyjmy się podstawowym (atomowym) elementom arkusza, którymi mogą być:

- Pole tekstowe (pojedynczy wiersz), patrz rys. 58; (w HTML element formularza określany jako TYPE="TEXT")



Rys. 58. Pole tekstowe

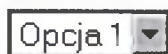
- Wielowierszowe pole tekstowe (obszar tekstowy), patrz rys. 59; (w HTML element formularza określany jako <TEXTAREA></TEXTAREA>)



Rys. 59. Wielowierszowe pole tekstowe

- Pole typu kombi, rys. 60 (w HTML element formularza określany jako

```
<SELECT>
<OPTION VALUE="Opcja 1">Opcja 1
<OPTION VALUE="Opcja 2">
...
</SELECT>
```



Rys. 60. Pole typu kombi pojedynczego wyboru

- Wielowierszowe pole kombi z możliwością zaznaczenia więcej niż jednego elementu, rys. 61 (w HTML element formularza określany jako

```
<SELECT MULTIPLE SIZE=4>
<OPTION VALUE="Opcja 1">Opcja 1
<OPTION VALUE="Opcja 2">
...
</SELECT>
```

⁶ Nazwa zbioru i ścieżka są zdefiniowane w zbiorze konfiguracyjnym. W przypadku standardowej instalacji serwera Xitami i aplikacji DEMO to prawdopodobnie będzie katalog: .../Xitami/Webpages/DEMO.



Rys. 61. Pole kombi wielokrotnego wyboru

- Przycisk opcji, pole typu „zaznacz” z możliwością wyboru więcej niż jednej opcji, rys. 62 (w HTML element formularza określany jako TYPE="CHECKBOX").



Rys. 62. Pole typu zaznacz (check)

- Przycisk opcji, pole typu „przełącznik” z możliwością wyboru jednej z wykluczających się opcji, rys. 63 (w HTML element formularza określany jako TYPE="RADIO").



Rys. 63. Pole typu przełącznik (radio)

Wymienione wyżej podstawowe elementy arkusza wprowadzania danych mogą być przyporządkowane do:

- a) pojedynczego pola bazy ISIS (powtarzalnego lub niepowtarzalnego),
- b) każdego podpola bazy ISIS (powtarzalnego lub niepowtarzalnego).

W przypadku (a) jeden element wprowadzania danych może być przyporządkowany do jednego pola ISIS (wraz z podpolami lub bez nich, powtarzalnego lub niepowtarzalnego). Zaś odpowiedzialność za właściwe wprowadzenie samych danych uwzględniające znaczniki podpól oraz wystąpienia pól (w WWW-ISIS jest to średnik), spoczywa przede wszystkim na końcowym użytkowniku.

W przypadku (b) możliwe jest rozbiecie pól bazy ISIS (ze zdefiniowanymi podpolami) na szereg elementów wprowadzania danych (pola typu kombi, lista, obszar tekstowy itd.). Czyli, jeśli docelowe pole bazy ISIS składa się z podpól $\wedge a$ oraz $\wedge b$, to możemy przyporządkować podpolu $\wedge a$ element tekstowy (pole tekstowe w typowym formularzu HTML), zaś w podpolu $\wedge b$ przycisk typu radio. Jeśli pole wraz z podpolami jest powtarzalne, to konieczne jest utworzenie osobnej grupy dla każdego wystąpienia (pola, podpola ISIS), czyli utworzenia tzw. grupy powtarzalnej. W takim przypadku formularz do wprowadzania danych będzie posiadał dodatkowy przycisk z napisem **ADD GROUP**, który pozwala na wyświetlanie kolejnych wystąpień danych elementów formularza, jako wystąpienie tego samego pola.

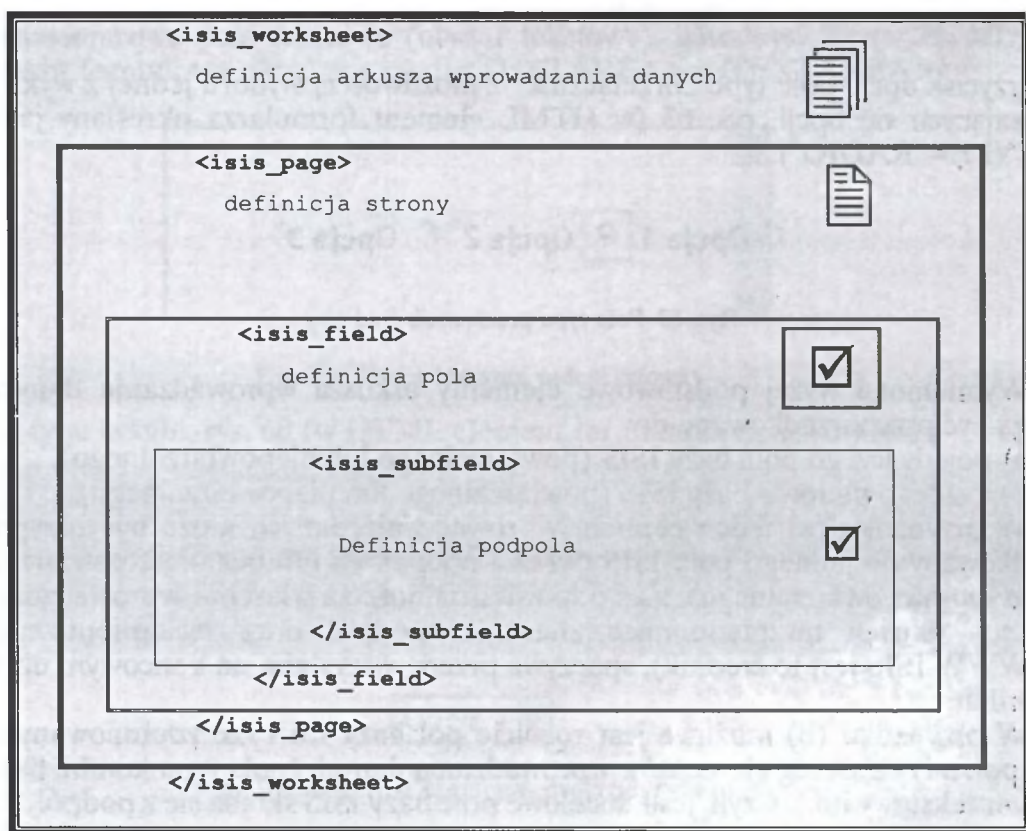
Tablica definicji arkuszy wprowadzania danych jest zapisana w zbiorze Entry.def. Zbiór ten podzielony jest na poszczególne bloki, związane z poszczególnymi typami rekordu. Każdy zaś blok (opisujący dany arkusz wprowadzania danych) składa się z opisu poszczególnych stron wchodzących w skład danego

rekordu. Sekcje dotyczące stron z kolei opisują elementy wprowadzania danych (takie jak wielowierszowe pole tekstowe, pole typu kombi, pole typu zaznacz itd.), a także specyfikacje innych elementów (tekst, elementy graficzne itd.). Zastosowano dość prosty język skryptowy a wszelkie konstrukcje mają strukturę hierarchiczną:

- tablica składa się z opisu poszczególnych arkuszy roboczych (worksheet),
- każdy arkusz roboczy składa się z definicji stron (page),
- każda strona opisuje poszczególne pola ISIS (field).

Ponadto każdy element opisujący pole ISIS może składać się z opisu podpól (jeśli występują w danej bazie ISIS) (subfield).

Konwencja zastosowana w opisie entry.def (tablicy definicji arkuszy wprowadzania danych) przypomina poprawny strukturalnie zbiór XML (każda sekcja posiada znacznik zamykający i otwierający), schematycznie całość przedstawia rys. 64



Rys. 64. Struktura zbioru entry.def (tablica definicji arkuszy wprowadzania danych)

Definicja arkusza wprowadzania danych składa się z definicji stron (jednej lub wielu). Zaś definicja strony zawiera definicje poszczególnych pól związanych z polami ISIS. Te ostatnie składają się z wielu parametrów i mogą dodatkowo zawierać sekcje podpól ISIS.

Definicja podpól składa się z odpowiednich parametrów. Poniżej zostaną przedstawione szczegółowe uwagi dotyczące składni każdego skryptu.

- Każda sekcja zaczyna się i kończy osobnym wierszem.
- Definicja każdego parametru kończy się średnikiem.

Poniżej przedstawiamy szczegółowo każdą z sekcji.

5.4.1. Sekcja arkusza wprowadzania danych



Głównymi elementami zbioru entry.def tablicy definicji arkuszy wprowadzania danych są definicje arkuszy umieszczane między znacznikami <isis_worksheet> oraz </isis_worksheet>.

W sekcji arkusza wprowadzania danych zdefiniowane są następujące parametry:

Tablica 4

Parametry arkusza wprowadzania danych (isis_worksheet)

Nazwa parametru	Status	Wyjaśnienie
Name	Obowiązkowy	Parametr ten podaje nazwę arkusza, używaną przez program. Na przykład połączenie wywołujące arkusz MON w celu wprowadzenia nowego rekordu może być określone jako: de_worksheet=MON
Title	Obowiązkowy	Parametr ten powoduje wyświetlenie tytułu arkusza (przeważnie w górnej części ramki HTML).
dbn	Obowiązkowy	Parametr ten definiuje bazę ISIS, do której odwołuje się dany arkusz.
Path	Opcjonalny	Parametr ten określa położenie (katalog), w którym przechowywane są bazy ISIS wraz z koniecznymi formatami oraz tablicami FST, FDT itd. Jeśli nie zostanie on określony, to system użyje domyślnie definicji w zbiorze konfiguracyjnym (DBNFile=...).
Valid_pft	Opcjonalny	Określa nazwę formatu walidacji danych.
Vdt	Opcjonalny	Określa nazwę tablicy walidacji danych.
Del_pft	Opcjonalny	Określa nazwę formatu kontrolującego usuwanie danych.
On_new_pft	Opcjonalny	Określa nazwę formatu uruchamiającego odpowiednie polecenia w momencie tworzenia nowego rekordu.
On_update_pft	Opcjonalny	Określa nazwę formatu uaktualniania rekordu.
Proof_pft	Opcjonalny, zalecany	Używany w celu wyświetlania rekordu zaraz po jego zapisaniu.
Magic	Zalecany	Parametr ten może być wykorzystywany do przechowywania numeru wersji arkusza wprowadzania danych. Jeśli w czasie korzystania z systemu okaże się, że użytkownik korzysta z wersji arkusza, która jest różna od tej określonej w tablicy definicji wprowadzania danych, to system może odrzucić wprowadzone do bazy rekordy lub poprawki. Parametr ten poprawia kontrolę dostępu do bazy osobom nieuprawnionym oraz ułatwia jej utrzymanie.

Przykład 10

Dla przykładowego arkusza „paper from journal” określone są następujące parametry:



```
name = ASEE;  
title = DEMO – AS (English);
```

```

dbn = demo;
path=;
vdt = EVASE.TAB;
valid_pft = EVASE;
del_pft = EDASE;
on_new_pft = @IASEE;
on_update_pft = @UASEE;
magic = 090320001102;

```

Jak więc widać, arkusz wprowadzania danych nazywa się ASEE, zaś w formularzu HTML wyświetlany jest tytuł „DEMO – AS (English)”. Baza danych, do jakiej odwołuje się arkusz, nazywa się demo i nie jest dla niej określona osobna ścieżka dostępu. Format kontroli na poziomie rekordu nazywa się EVASE.PFT, zaś tablica walidacyjna znajduje się w zbiorze EVASE.TAB, Ponadto format kontrolujący usuwanie rekordów znajduje się w zbiorze EDASE.PFT.

5.4.2. Sekcja stron

W obrębie sekcji arkusza wprowadzania danych (worksheet) wprowadzone są definicje poszczególnych stron umieszczane między znacznikami: `<isis_page>...</isis_page>`. Definicja każdej strony (`isis_page`) w tablicy definicji arkuszy wprowadzania danych rozpoczyna się definicją określającą nazwę (Name) i tytuł strony wyświetlany na formularzu HTML (Title). Ponadto konieczne jest podanie parametru – **Number** określającego kolejny numer strony.



Przykład 11

```

Name = MainPage;
Number = 1;
Title = Main bibliographic data;

```



W naszym przykładzie strona nazywa się MainPage, jej kolejny numer wynosi 1, a tytuł wyświetlany w formularzu HTML to „Main bibliographic data”.

Po podaniu głównych parametrów strony należy podać parametry określające przyporządkowanie pól ISIS, czyli sekcję pól.

5.4.3. Sekcja pól ISIS

W obrębie sekcji definiującej strony (`<isis_page>`), umieszczane są definicje pól pomiędzy znacznikami `<isis_field>...</isis_field>`.

Sekcja pól ISIS (`isis_field`) definiuje jeden lub więcej elementów wprowadzania danych, przyporządkowując je polom bazy ISIS.

Ze względu na specyfikę baz danych ISIS dopuszcza się dwojaki sposób rozwiązania:

1. Pole bazy ISIS może być przyporządkowane jednemu elementowi wprowadzania danych (takie rozwiązanie jest zalecane w przypadku, kiedy pole nie zawiera żadnych podpól). Wówczas składnia będzie miała następującą postać:

```
<isis_field>
  name = niepowtarzalna_nazwa_na_uzytek_cgi;
  description = opis_widoczny_na_formularzu; tag=nnn;
  [definicja_elementu_formularza_wprowadzania_danych]
</isis_field>
```

2. Pole bazy ISIS składa się z podpól. W takim przypadku zaleca się, aby każdemu podpolu bazy ISIS przyporządkować jeden element wprowadzania danych, zaś składnia będzie miała następującą postać:

```
<isis_field>
  name = niepowtarzalna_nazwa_na_uzytek_cgi;
  description= opis_widoczny_na_formularzu; tag=nnn;
  [sekcja_podpol_isis1;... sekcja_podpol_isisn]
</isis_field>
```

W każdym z wymienionych wyżej przypadków sekcja definicji pól ISIS zawiera następujące wspólne parametry:

Name – określa łatwą do zapamiętania nazwę, używaną przez program cgi do identyfikacji elementu i przetwarzania danych (nazwy oczywiście nie mogą się powtarzać).

Description – opis pola pojawiający się z arkusza.

Tag – parametr przyporządkujący znacznik ISIS do pola.

DLA PÓL Z PODPOLAMI:

Jeśli pole bazy ISIS składa się z podpól, to zaleca się, aby było ono reprezentowane na arkuszu przez kilka elementów wprowadzania danych i każdy z nich był przyporządkowany do odpowiedniego podpola.



DLA PÓL (PODPÓL) POWTARZALNYCH:

Jeśli pole bazy ISIS ma podpola i jest powtarzalne, to powinniśmy zastosować tzw. grupy powtarzalne chyba że dane te są automatycznie uzupełnianie przez tabele odnośników (por. przykł. 17 oraz przykł. 18)



Odpowiednio w zależności od sytuacji (1 lub 2) możemy mieć następujące przypadki:

Przykład 12

Przykład ten ilustruje przypadek 1 (pole bazy ISIS przyporządkowane jednemu elementowi wprowadzania danych). Pokazana jest definicja pola nazwanego TRN na arkuszu.



```
<isis_field>
  name = trn; description = TRN (1); tag=1;
  type=1; size = 13; max_size = 13;
</isis_field>
```

Jak widać znacznikowi 1 reprezentującemu TRN (Temporary Record Number) przyporządkowany jest jeden element. Element ten posiada typ 1 (czyli na formularzu HTML będzie widoczne jednowierszowe pole tekstowe), jego rozmiar będzie wynosił 13 i będzie można wprowadzić maksymalnie 13 znaków (max_size = 13).

5.4.4. Sekcja podpól

Składnia sekcji podpól jest bardzo podobna do przypadku 1 (sekcji pól) tj.:

```
<isis_subfield>
  name = niepowtarzalna_nazwa_na_uzytek_cgi;
  description= opis_widoczny_na_formularzu; tag=nnn;
  tag=^c;
  [definicja_elementu_formularza_wprowadzania_danych]
</isis_subfield>
```

Jak widać, jedyną różnicą jest wartość parametru określającego znacznik pola i z oczywistych powodów może przyjmować jako wartości znaczniki podpól z tablicy FDT.

Składnia dla części **[definicja_elementu_formularza_wprowadzania_danych]** jest oczywiście taka sama w przypadku pól, jak i podpól. Poniżej zostanie przedstawiona ona szczegółowo.

Podstawowe parametry wyglądają następująco:

Type – ten parametr jest obowiązkowy, opisuje typ elementu formularza wprowadzania danych. Może on mieć wartości podobne jak dla pola, zatem dopuszczalne są następujące możliwości:

Tablica 5

Typy elementów wprowadzania danych

type=	Opis	Przykładowy wygląd w formularzu HTML
1	Jednowierszowe pole tekstowe	
2	Wielowierszowe pole tekstowe ()	
3	Pole typu kombi pojedynczego wyboru	Opcja 1
4	Pole typu kombi wielokrotnego wyboru	Opcja 1 Opcja 2 Opcja 3 Opcja 4
5	Pole typu zaznacz (CHECKBOX)	☒ Opcja 1 ☒ Opcja 2 ☐ Opcja 3
6	Pole typu przełącznik (RADIO)	☐ Opcja 1 ☒ Opcja 2 ☐ Opcja 3
7	Komentarz	Komentarz
8	Pole chronione (bez możliwości wpisywania danych, np. do wyświetlania zawartości pól uzupełnianych automatycznie)	Bez możliwości wprowadzania danych

- default** – ten parametr jest opcjonalny; dostarcza wartości domyślnej w momencie wprowadzania danych do formularza, może być zastosowany dla dowolnego typu pola.
- repeat** – parametr ten stosowany w przypadku grup powtarzalnych określa, czy nowe wystąpienie danego pola powinno pojawić się na ekranie, dopuszczalne są następujące opcje (tab. 13).

Tabela 13

Parametry dla pól powtarzalnych

repeat=	Opis
Yes	W formularzu pojawia się specjalny przycisk ADD NEW wyświetlający następne wystąpienie danego pola i pozwalający użytkownikowi na wprowadzenie kolejnych danych.
No	Przycisk ADD NEW nie pojawi się na formularzu.

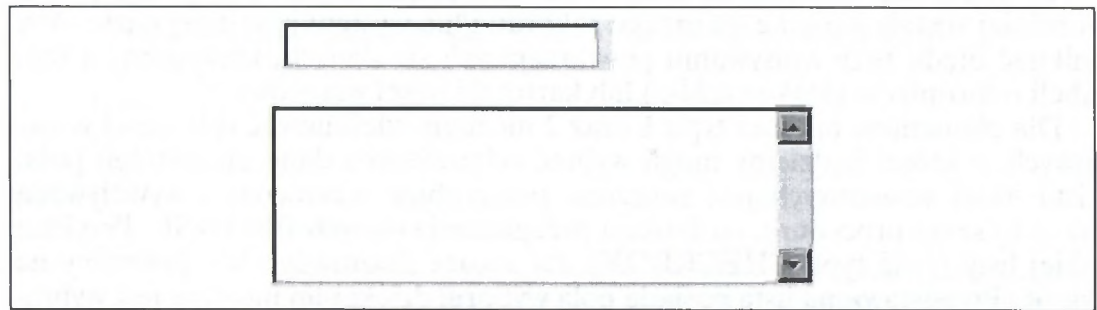
UWAGA:

Parametr repeat nie może być umieszczany w opisie sekcji dotyczących podpól⁷.

Dla specyficznych typów elementów wprowadzania danych należy określić dodatkowe parametry. Poniżej zdefiniujemy parametry charakterystyczne dla poszczególnych typów.

5.4.5. Definiowanie jedno i wielowierszowych pól tekstowych (typ 1 i 2)

Zajmiemy się szczegółowo parametrami definiowania pól jedno i wielowierszowych w sekcji pól (podpól) w tablicy definicji arkuszy wprowadzania danych.



Rys. 65. Jedno i wielowierszowe pola tekstowe (typ: 1, 2)

- Dla przypomnienia w formularzach HTML:
- pola jednowierszowe definiowane są jako `<INPUT TYPE="TEXT";`
 - zaś pola wielowierszowe (obszar tekstowy) definiowane są jako `<TEXTAREA> </TEXTAREA>;`

⁷ Nie spodziewamy się podpól powtarzalnych, bo nie jest to zgodne z językiem formatowania ISIS.

Size – parametr określa długość pola tekstowego *text* oraz obszaru tekstowego *text-area*; długość możemy określić w dwojaki sposób: **small, medium, large** (odpowiednio mała/y, średni/a, duży/a) lub **jako wartości liczbowe** (przez liczby naturalne – N). Jeśli element jest jednowierszowym polem tekstowym (**type=1**), N oznacza liczbę znaków w polu (przesuwalnym); jeśli element jest wielowierszowym polem tekstowym (**type=2**), wymagane są dwa parametry: liczba wierszy oraz liczba kolumn (N);

Rows – patrz wyżej (istotne tylko w przypadku wielowierszowych pól tekstowych – *text-area*);

Max_size – maksymalna długość tekstu, który może być wprowadzany do elementu arkusza.



UWAGA:

Parametry *size*, *rows*, *max_size* będą zignorowane jeśli wybierzemy typy 3-6.



Przykład 13

Aby zdefiniować element arkusza typu tekst musimy napisać:

Type = 1; Size = medium

Aby zdefiniować obszar tekstowy (50 kolumn, 5 wierszy) zapisujemy:

Type = 2; Size = 50; rows=5;

Aby zdefiniować pole tekstowe dla wprowadzania roku należy wpisać:

Type = 1; Size = 4; max_size=4;



5.4.6. Parametry odsyłaczy

W momencie wpisywania danych do formularza często zdarza się, że już wcześniej zostały wpisane do innego rekordu i/lub występują w innej bazie. Aby uniknąć błędu przy wpisywaniu powtarzających się danych, korzystamy z tzw. tabeli odnośników (*lookup tables*) lub kartoteki haseł wzorcowych.

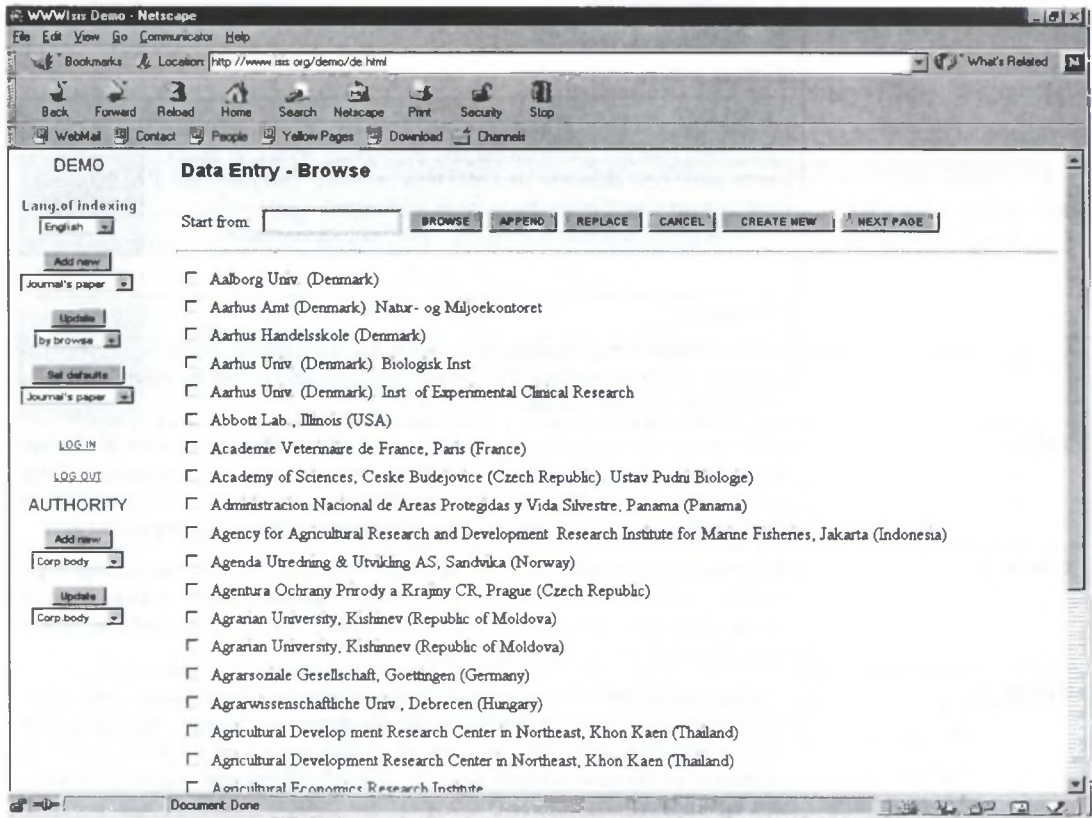
Dla elementów arkusza typu 1 oraz 2 możemy zdefiniować listę haseł wzorcowych, z której będziemy mogli wybrać odpowiednie dane do naszego pola. Lista haseł wzorcowych jest tworzona przez zbiór odwrócony i wywoływana przez tę samą procedurę, co funkcja przeglądania danych BROWSE. Przykład takiej listy (pola typu CHECKBOX) dla autora zbiorowego jest pokazany na rys. 66. Przedstawiona lista posiada pola wyboru, dzięki nim możliwe jest wybranie więcej niż jednego autora zbiorowego do naszego pola. Możliwe jest również stworzenie podobnej listy odsyłaczy, lecz z przyciskiem opcji (czyli możliwa jest tylko jedna wartość do zaznaczenia). Podobnie, jak w przypadku funkcji BROWSE (przeglądania danych), aby możliwe było wyświetlenie takiej listy, konieczne jest ustalenie:

1. Prefiksów, dzięki którym tablica taka jest wyświetlana.

2. Formatu, dzięki któremu można wyświetlać poszczególne wiersze.

3. Może się zdarzyć, że będziemy chcieli wyświetlać co innego niż wartość, która powinna zostać skopiowana do pola. Na przykład, gdy wartości pola są kodowane,

wygodniej jest oglądać tabelę z wartościami rozwiniętymi, zaś w chwili wyboru ściągnąć wartość zakodowaną. W takim przypadku należy stworzyć odpowiedni format, w którym wyświetlamy rozwinięcia, zaś jako wartość selekcji określamy kod.



Rys. 66. Przykład listy odsyłaczy (pół wyboru)

Bardziej wyczerpująca lista parametrów dla tablic odnośników przedstawiona jest w tablicy 6.

Tablica 6
Parametry dla tablic odnośników (interfejs wprowadzania danych)

Parametr	Jego działanie
1	2
Dbn	Nazwa bazy , jeśli brak, system zakłada, że domyślną bazą jest ta, z którą aktualnie pracuje użytkownik.
Path	Ścieżka dostępu do baz (opcjonalnie). Jeśli nie podamy żadnej ścieżki (zalecane), ścieżka pobierana jest ze zbiorów określających bazę danych (dbn.def)
Pfx	Prefiks wskazuje, którą część zbioru odwróconego należy wykorzystać.
Fmt	Wykorzystywany format .
Startterm	Termin (ciąg znaków), od którego (w kolejności alfabetycznej) zaczyna się wyświetlanie listy odsyłaczy, domyślnie jest pusty.
lookup_type	Ten parametr określa, czy istnieje tablica odsyłaczy czy nie. Jeśli wartość jest równa 0, to nie zostanie utworzona lista odsyłaczy; jeśli wartość będzie wynosić 2 to pojawi się przycisk BROWSE .

1	2
Lookup_choice	Parametry te określają, czy będzie możliwe zaznaczanie więcej niż jednej opcji z listy wyboru. Możliwe są następujące wartości: 1 – pojedynczy wybór 2 – wielokrotny wybór 3 – brak wyboru UWAGA: Jeśli dana lista jest zdefiniowana jako pojedynczy wybór, to zaleca się ustawienie typu elementu jako pole typu „zaznacz” (przycisk typu RADIO); jeśli chcemy mieć możliwość zaznaczania wielu opcji – pole wyboru (CHECKBOX), w przeciwnym wypadku nie będzie wyświetlona żadna opcja. UWAGA: Pola bez możliwości wyboru mogą okazać się przydatne dla pól ISIS, których wartości powinny być niepowtarzalne, jak np. numer rekordu, ISSN, Numer ID użytkownika itd.
Table	Jeśli zamierzamy umieścić możliwość przeglądania danych zgodnie z różnymi kryteriami sortowania danych, najlepiej odwołać się do stosownej tablicy przeglądania danych BROWSE.DEF (w interfejsie wyszukiwania danych WWW-ISIS), wartością pola jest więc nazwa stosownej tablicy wyszukiwań .
Dlmtr_l	Jeśli zostało zastosowane pole wielokrotnego wyboru (możliwość zaznaczenia wielu wartości jednocześnie) parametr ten określa lewostronny ogranicznik pola (separator), który będzie dodany do każdej pozycji po jej lewej stronie w momencie zapisywania rekordu do bazy.
Dlmtr_r	Jeśli zostało zastosowane pole wielokrotnego wyboru (możliwość zaznaczenia wielu wartości jednocześnie) parametr ten określa prawostronny ogranicznik pola (separator), który będzie dodany do każdej pozycji po jej prawej stronie w momencie zapisywania rekordu do bazy. Brak określenia tego parametru oznacza brak ogranicznika. Przykład: dla pola deskryptorów można zdefiniować lewe i prawe ograniczniki jak następuje: dlmtr_l = < ; dlmtr_r = > ; Jeśli zostaną wybrane terminy ABA oraz LAND to zostaną one wprowadzone do bazy jako: <ABA><LAND>
Gen_occ	Możliwe wartości to: YES, NO . Jeśli zostało zastosowane pole wielokrotnego wyboru, parametr ten określa, że zaznaczony fragment z listy będzie wprowadzony do bazy jako nowe (kolejne) wystąpienie danego pola.
Create	Opcja ta umożliwia końcowym użytkownikom dodanie bazy haseł wzorcowych, kiedy nie jest możliwe odnalezienie odpowiednich danych. Możliwe wartości to YES, NO (domyślne jest NO). Jeśli wprowadzony zostanie ten parametr (wartość YES), na ekranie z hasłami wzorcowymi pojawi się przycisk CREATE NEW. Zastosowanie tej funkcji pozwala utworzyć brakujące hasło i powrócić do tworzenia lub aktualizacji rekordu, dla którego to hasło było potrzebne.
Create_type	Parametr ten może być użyty wraz z opcją <i>create</i> . Wskazuje on na nazwę arkusza , która powinna być podana ilekroć końcowy użytkownik bazy zamierza wprowadzić brakujący element.
Init_pft	Parametr ten podaje nazwę formatu , który inicjuje wartości początkowe nowego rekordu (po przyciśnięciu przycisku CREATE NEW w ekranie przeglądania danych (por. parametr <i>create</i>)).

Poniżej zilustrujemy wykorzystanie parametrów z tabeli opisującej parametry dla tablic odnośników (interfejs wprowadzania danych).

Przykład 14

Załóżmy, że zamierzamy zdefiniować element arkusza wprowadzania danych dla pola „abstrakty” (w bazie ISIS jest to pole powtarzalne). Definicja takiego elementu wyglądałaby następująco:



```
<isis_field>
  name = abstract; description = Abstract (714); tag=714; type=2;
  size=50; rows = 5; lookup_type=0;
</isis_field>
```

Tak więc w naszej specyfikacji definiujemy obszar tekstowy dla pola bazy ISIS o znaczniku 714. Rozmiar naszego obszaru tekstowego wynosi wysokość – 5 wierszy oraz długość 50 znaków w wierszu.

Rys. 67. Przykład wielowierszowego pola tekstowego

Przykład 15

Załóżmy, że chcemy zdefiniować element wprowadzania danych dla autorów zbiorowych (w bazie ISIS jest to pole powtarzalne). Chcielibyśmy dołączyć opcje wyboru danych z tablicy odnośników (*lookup table*) zbudowanej z bazy A-CORP. Ponadto jeśli użytkownik nie znajdzie we wspomnianej bazie odnośników interesujących go danych, powinien być w stanie dodać element do bazy odnośników A-CORP. Definicja wyglądałaby następująco:



```
<isis_field>
  name = cauthor; description = Corporate author (110);
  tag=110; type=1; size=50;
  lookup_type=2; lookup_choice=2; pfx=EA;; dbn = AUTH; fmt=EBCAR;
  gen_occ=yes; create = yes; create_type = COR;
</isis_field>
```

W tej definicji element (pole tekstowe) arkusza odpowiada polu 110 w bazie ISIS.

Rys. 68. Pole "Corporate author" z tablicą odnośników

Definiując tablicę odnośników wprowadziliśmy przycisk **BROWSE** (lookup_type=2). Tablica odnośników daje możliwość wielokrotnego wyboru (lookup_choice=2). Jest to określone w następujący sposób: wykorzystujemy technologię zbioru odwróconego z bazy AUTH; część zaczyna się prefiksem EA: (pfx=EA;; dbn = AUTH;). Dla każdego odwołania do punktu dostępu do wspomnianej bazy używamy prefiksu "EA:", zastosowany jest format EBCAR.PFT.



UWAGA:

To, w jaki sposób zostaną wyświetlone elementy selekcji danych z bazy, zależy od formatu!

Wartości, które można przeglądać, są podzielone na strony (por. parametr stronicowania wyników w zbiorze konfiguracyjnym, Aneks A).

Ponadto ekran przeglądania danych posiada przycisk **CREATE NEW** pozwalający na tworzenie nowego rekordu dla tablicy odnośników (baza AUTH – Corporate Author), jeśli nie znajduje się w bazie.



WSKAZÓWKA:

Kiedy końcowy użytkownik odwołuje się do bazy odnośników i dostępny jest przycisk **CREATE NEW**, po jego przyciśnięciu możliwe jest wprowadzenie danych w określonym arkuszu, a wskazane wartości z tablicy odnośników będą skopiowane do danego pola.



UWAGA:

Procedura łączenia rekordów może mieć charakter rekursywny tzn. z rekordu A możemy odwoływać się do listy odnośników, a jeśli zachodzi taka konieczność możemy stworzyć rekord B i wywołać inną tablicę odnośników, a następnie zacząć tworzyć rekord C itd., aby na końcu powrócić do procesu tworzenia rekordu A i zakończyć go (podobnie z rekordem B itd). Z uwagi na ograniczenia o charakterze ergonomicznym nie zaleca się tworzenia więcej niż dwóch poziomów w funkcji CREATE.

Wybór jednego pola

Wybór listy odnośników dla pola Corporate Author (autor zbiorowy) jest tworzony przez odpowiedni format (w tym konkretnym przypadku to EB_CAR). Pod warunkiem, że tablica odnośników zawiera nazwę autora w polu v120, format może wyglądać następująco:

```
'<TR><TD VALIGN=TOP><INPUT TYPE="RADIO" NAME="de_browse_name"
VALUE=" 'v120' ">' '<TD>'v120'</TR>'
```

Przeanalizujemy powyższy format bardziej szczegółowo. Występuje w nim wiersz tworzący element formularza HTML tzw. przycisk typu „przełącznik” RADIO (), o nazwie *de_browse_name* i przyjmujący wartość z pola o znaczniku 120, w konsekwencji będzie wyświetlany tekst wyjaśniający (w tym przypadku zawartość pola v120). Technika ta pozwala na wyprowadzanie tekstu wyjaśniającego różnego od zaznaczonej wartości. Ponadto pozwala na wyświetlanie elementów danych dłuższych niż terminy w zbiorze odwróconym.

Wybór grup pól

Może się zdarzyć, że chcielibyśmy „zaimportować” więcej niż jedno pole z rekordu wskazanego na liście odnośników. Aby lepiej wyjaśnić tę technikę przyjrzymy się następującemu eksperymentowi:

1. Najpierw wprowadźmy nowy rekord dot. czasopism (Authority part).

Serial title	Intelligent Information Systems	BROWSE
ISSN	0925-9902	BROWSE
Availability	For the staff only	
Online availability		

Rys. 69. Wprowadzanie nowego rekordu do tabeli odnośników

2. Następnie zacznijmy wprowadzać rekord o typie AS (formularz AS) i na stronie 4 wywołajmy listę tytułów czasopism a następnie wybierzmy nowo wprowadzony tytuł.

3. Po powrocie do rekordu dot. czasopisma zauważymy, że nie tylko tytuł został wybrany, ale też zostały zaimportowane stosowne dane, m.in. ISSN oraz informacje na temat dostępności (Availability) (rys. 69).

Aby zaimportować pewne wartości ze wskazanego rekordu do innych pól w formularzu format winien posiadać następującą strukturę:

```
<INPUT TYPE="RADIO" NAME="de_browse_name"
VALUE="selection data###tt1### wartość dla pola tt1### tt2
wartość dla pola tt2 etc.
```

gdzie tt₁ jest znacznikiem w arkuszu wprowadzania danych, do którego będą wprowadzone dane wartość dla pola tt₁.

Następne dwa przykłady ilustrują sposób, w jaki możemy zaprojektować format do tworzenia tablicy wyboru wzorców w celu skopiowania więcej niż jednej wartości.

Przykład 16

Zalóżmy, że w rekordach dotyczących czasopism numer ISSN jest w polu v21, zaś dostępność (availability) w polu V483. Możemy zatem zaprojektować następujący format przeglądania danych:



```
'<TR><TD VALIGN=TOP><INPUT TYPE="RADIO" NAME="de_browse_name"
VALUE="'
v421,
"###450#####"v27,
"###483#####"v483,
'"><TD>',
v421,"; ISSN "v27,'</TR>'
```

Dzięki temu formatowi możemy wyświetlić w każdym wierszu przycisk opcji (radio), po którym następuje zawartość pola v421 (serial title) oraz ISSN. Ponadto możemy określić parametr VALUE przez zawartość pola v421, po któ-

rym zostaną umieszczone dane do zaimportowania pola 450 (z pola ISSN – v27 rekordów dot. czasopism) oraz pola 483 (z pola availability – v483 rekord czasopism – Record)



Przykład 17

Załóżmy, że w arkuszu mamy pozycję (pole) "Corporate Author", po której jest pozycja "Address" przyporządkowana do pola 311. Baza wzorcowa rekordów Autorów Zbiorowych ("Corporate Author") zawiera nazwę w polu v120 oraz „Address” w polu v121. Jeśli chcemy, aby wraz z wyborem danych z listy nie tylko była pobierana nazwa autora zbiorowego ale także jego adres (Address), dla pola v311 powinniśmy zaprojektować następujący format:

```
'<TR><TD VALIGN=TOP><INPUT TYPE="RADIO" NAME="de_browse_name"
VALUE=" ',v120,'###311###',v121,' ">'
'<TD>'v120'</TR>'
```

Jak widać parametr „value” będzie zawierał adres pobrany z pola 121 zaś oznaczony będzie znacznikiem 311, tak więc system wprowadzi wartość do pola o znaczniku v311. Jeśli zawiera on już jakąś wartość to będzie ona zamieniona na nowo wprowadzoną.



REGUŁY KOPIOWANIA DANYCH:

1. Można kopiować pola niepowtarzalne do pól niepowtarzalnych; np. tytuł czasopisma wraz z numerem ISSN.
2. Jeśli docelowym elementem jest grupa powtarzalna, możliwe jest zaprojektowanie takiego formatu, który będzie „importował” dane z każdej powtarzalnej grupy. W obrębie jednej grupy żadne pole nie może być wywołane więcej niż raz; jeśli projektujemy format dla grup powtarzalnych, żadne dane spoza grupy nie mogą być importowane.
3. Jeśli ze wskazanego rekordu chcielibyśmy skopiować powtarzalne pole (np. wybranych dwóch lub trzech autorów), w arkuszu wprowadzania danych nie możemy utworzyć grupy powtarzalnej, zamiast tego zaleca się zaimportowanie wszystkich powtarzalnych danych do jednego pola (pola tekstowego lub obszaru tekstowego), dane powinny być oddzielone średnikiem.

Następny przykład ilustruje aplikacje, w których zastosowano ograniczniki „<” oraz „>” wykorzystywane w celu wprowadzania danych w nawiasach trójkątnych. Ponadto pokażemy, jak zdefiniować zaawansowane tablice odnośników na podstawie tezauryśa dołączane do pola deskryptorów.

Przykład 18



Załóżmy, że chcielibyśmy zaimplementować opcję wyboru deskryptorów z bazy tezauryśa nazwanej THES (rys. 70). Użytkownik powinien być w stanie wprowadzić ręcznie deskryptory, rozdzielając je średnikami lub też móc wybierać je z tablicy odnośników (*lookup table*). Zakładamy, że dodawanie dowolnego terminu do tezauryśa powinno być zablokowane. Nasza definicja w tablicy definicji arkuszy wprowadzania danych wyglądałaby następująco:

```
<isis_field>
name = descr; title = Descriptors; tag= 501;
type=1; size=medium; lookup_type=2; lookup_choice=2;
dbn=THES;
pfx = E;; pft= DSC-SEL; table=DSC1;
dlmtr_l = <; dlmtr_r=>;
</isis_field>
```

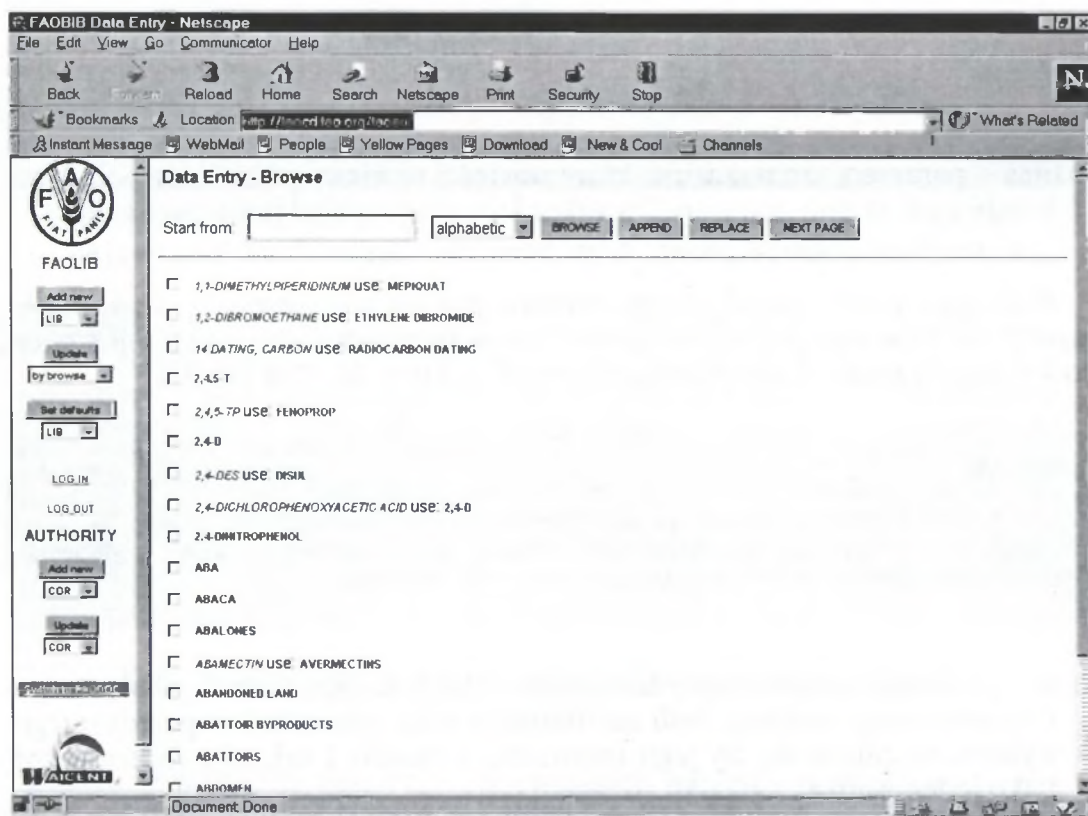


Definiujemy wielowierszowe pola tekstowe dla niepowtarzalnego pola o znaczniku v521 (w bazie ISIS). Wprowadzamy przycisk LIST. Dzięki niemu możliwe będzie wyświetlenie listy z polami typu CHECKBOX (pozwalającymi na zaznaczenie wielu pól jednocześnie). Ponieważ odwołujemy się do tablicy DSC1, wyświetlane w formularzu pole typu kombi pojedynczego wyboru umożliwia przeglądanie tezaursu na różne sposoby: deskryptory w porządku alfabetycznym lub jako tzw. KWOC.

Kiedy będziemy wybierali terminy z listy, pojawią się one w obszarze tekstowym oddzielone od siebie średnikiem. Kiedy zapiszemy je do bazy zostaną one ujęte w nawiasy trójkątne, zdefiniowane przez nas jako lewy i prawy ogranicznik („<” oraz „>”)

Poniżej przedstawiony jest pełen format pozwalający na wybór danych:

```
'<TR><TD VALIGN=TOP><INPUT TYPE="CHECKBOX" NAME="de_browse_name" VALUE=" '
if p(v001) then v001 else V003 fi,
'"><TD>'
if p(v001) then
'<FONT FACE="Arial,Helvetica" COLOR=GREEN SIZE=2><B>'v001'</B></FONT> ',
else
'<FONT FACE="Arial,Helvetica" COLOR=RED SIZE=2><I>'v008'</I></FONT> '
,D008" use: ",
'<FONT FACE="Arial,Helvetica" COLOR=GREEN SIZE=2><B>'v003'</B></FONT> ',
fi,
'</TR>'
```



Rys. 70. Lista z polami typu „zaznacz” w tezaursie

Przedstawiony format wyświetla każdy wiersz zaczynając od pola wyboru. Askryptory (pole v8) na ekranie komputera są zaznaczane kolorem czerwonym.

W przypadku wielojęzycznego tezaursu można bez problemu zaimplementować listę, w której deskryptory wyświetlone są w języku angielskim, lecz wybierane są ich francuskie odpowiedniki lub na odwrót.



UWAGA:

Jeśli chcielibyśmy wyświetlać wartości ze zbioru odwróconego, wartość „term” (terminy ze słownika) jest przechowywana w wirtualnym polu o znaczniku 9980.

5.4.7. Inne elementy formularza wprowadzania danych

Pozostało jeszcze kilka innych składników formularza, jak: pola wyboru, pola kombi itd. Poniżej zostaną omówione szczegółowo parametry konieczne do ich zdefiniowania.

Parametry wspólne dla pól kombi pojedynczego i wielokrotnego wyboru (typ 3 i 4).



Rys. 71. Pola typu kombi pojedynczego i wielokrotnego wyboru (typ 3 i 4)

Values – parametr ten wskazuje, które wartości powinny pojawić się w polu typu kombi i jak są one „zapisane” w rekordzie. Na przykład jeżeli zapiszemy:

`values=Monograph|M,paper from book|AM, paper from journal|AS`

Pole typu kombi pojedynczego wyboru pokaże trzy wartości: *Monograph*, *paper from book* oraz *paper from journal*, zaś w przypadku wyboru którejs z opcji do każdego rekordu dodane będą odpowiednio kody: M, AM lub AS.



UWAGA:

Jeśli zdefiniujemy pole, którego wypełnienie nie jest obowiązkowe, zaleca się, aby pierwsza wartość była pustym łańcuchem znaków. W ten sposób końcowy użytkownik będzie miał możliwość niewprowadzania wartości do tego pola.

Size – parametr ten definiuje rozmiar okna dla pola typu kombi pojedynczego i wielokrotnego wyboru. Jeśli zdefiniujemy pole typu kombi – pojedynczego wyboru, to zaleca się, by jego ustawienie wynosiło 1 tak, aby widoczny był tylko jeden wiersz.

Gen_occ – ten parametr może być zastosowany tylko w polach typu kombi – wielokrotnego wyboru. Znaczenie tego parametru jest analogiczne do pól typu

kombi pojedynczego wyboru, tzn. dla każdej wybranej wartości tworzone jest kolejne wystąpienie pola ISIS (dot. pól powtarzalnych).

Parametry wspólne dla pól typu „zaznacz” i „przełącznik” (typ 5 i 6).

☒ Opcja 1 ☒ Opcja 2 ☐ Opcja 3

☐ Opcja 1 ☒ Opcja 2 ☐ Opcja 3

Rys. 72. Pola typu „zaznacz” i „przełącznik” (typ 5 i 6)

Values – jak wyżej.

Gen_occ – parametr ten może być wybrany tylko dla pól wyboru (zaznaczanie wielu opcji); znaczenie analogiczne jak dla pól kombi.

Cols – liczba kolumn dla wartości na ekranie wprowadzania danych; dla przycisków typu radio zdefiniowane są dwie kolumny.

UWAGA 1:

Pole typu przełącznik (przycisk typu radio) powinno być użyte jedynie wtedy, gdy możliwe jest dokonanie pojedynczego wyboru, podczas gdy pola wyboru (przycisk typu zaznacz) – gdy użytkownik końcowy miałby wprowadzać więcej niż jedną wartość do danego pola.



UWAGA 2:

W przypadku pól typu przełącznik, jeśli pole ISIS (lub podpole) nie jest obowiązkowe powinniśmy dodatkowo wprowadzić opcję "Bez wyboru" z pustą wartością ("").



UWAGA 3:

W przypadku pola typu zaznacz należy odpowiednio zaprojektować jego (ich) układ i powiązanie z polami ISIS. Jeśli zbiór pól wyboru odpowiada niepowtarzalnemu polu ISIS, należy zdefiniować prawy i lewy separator np. „<”, „>”. Jeśli zaś zbiór pól wyboru odpowiada polu powtarzalnemu, należy dodać w definicji pól wyboru `gen_occ=yes` tak, aby każdy wybór pola tworzył nowe wystąpienie pola ISIS.



5.5. Narzędzia sprawdzania poprawności

System WWW-ISIS dostarcza odpowiednich narzędzi pozwalających definiować walidację wprowadzanych rekordów, czyli sprawdzać, czy dane spełniają nałożone na nie warunki poprawności. Zostaną omówione następujące struktury walidacyjne:

- tablice sprawdzania poprawności pól (kontrola zawartości pól względem zbiorów wartości zawartych w bazie ISIS);
- formaty globalnej poprawności (kontrola na poziomie rekordów);
- formaty nadzorujące usuwanie rekordów.

Formaty walidacyjne określone są w tablicy arkusza wprowadzania danych (patrz parametry: Valid_ptf, Del_pft, oraz Vdt).

5.5.1. Tablica poprawności pól

W każdym arkuszu wprowadzania danych można zdefiniować tablicę poprawności wprowadzanych danych dla pól, określając reguły sprawdzające spójność wprowadzanych do bazy danych (na poziomie pola). Tablica składa się z dowolnej liczby wierszy, zaś każdy z nich składa się z następujących elementów:

1. Element bazy danych ISIS, gdzie znajdują się wzorce danych.
2. Technika selekcji.
3. Tryb walidacji.
4. Specyfikacja poprawności.
5. Format komunikatu (w przypadku błędu).
6. Specyfikacja transformacji (opcjonalna).

Wszystkie wymienione elementy są obowiązkowe poza specyfikacją transformacji. Zanim rozpoczniemy analizę parametrów wyjaśnimy, jak system wykorzystuje tablicę. Dla każdego rekordu przed zapisem do bazy danych sprawdzany jest każdy wiersz tablicy w następujący sposób:

- W zależności od techniki selekcji wybierany jest element danych z przetwarzanego fragmentu danych ISIS (każde wystąpienie pola lub każdy deskryptor itd.).
- Wykonywane jest odpowiednie sprawdzanie danych względem bazy haseł wzorcowych (odnośników). Choć walidacja przeprowadzana jest względem zawartości zbioru odwróconego bazy haseł wzorcowych, to w przypadku haseł dłuższych niż 30 znaków system odwołuje się do oryginalnych wartości rekordów, których adresy wskazane są w zbiorze odwróconym.
- Jeśli wprowadzone dane będą niezgodne z regułami poprawności, to wirtualne pole v9999 jest wykorzystywane w celu przechowywania niezapisanych do bazy danych. Pole v9999 jest powtarzalne, tak więc jeśli w procesie wprowadzania danych wystąpi więcej niż jeden błąd, to wszystkie one są sekwencyjnie przechowywane. Zanim następny wiersz FST zostanie przetworzony, system wyświetla wiadomość zgodną z wartością parametru *message format in case of fail*.

Prześledźmy teraz szczegółowo poszczególne elementy.

1. Element bazy danych ISIS ma następującą postać

tag = znacznik pola (podpola)

i wskazuje, które pole bazy ISIS ma być poddane sprawdzaniu.

Przykład 19



Tag= V20

Tag = V30^a

Tag = (V45^b) (dla pól powtarzalnych)



2. Technika selekcji ma postać:

```
Select = n
```

gdzie **n** jest numerycznym kodem oznaczającym technikę selekcji (por. FST). W chwili obecnej dostępne są następujące techniki selekcji:

- 0 – pobierane jest każde wystąpienie pola i zastosowana walidacja określona przez parametry: valid mode oraz valid-specification
- 4 – pobierane jest każde słowo z pola i stosowana jest walidacja (do zastosowania w następnej wersji systemu WWW-ISIS).

3. Tryb walidacji określa dostępne tryby poprawności; struktura jest następująca:

```
vmode = {AF, int, real, date} ; term = opcja1;
```

gdzie

- **AF** – opcja ta jest wykorzystywana, jeśli walidacja wykorzystuje bazę wzorców⁸;
- **Int** – jeśli element danych jest liczbą naturalną;
- **Real** – jeśli element danych jest liczbą rzeczywistą;
- **Date** – jeśli element danych jest daną typu data.

W przypadku zastosowania vmode=AF, opcja 1 może przyjąć następującą postać:

Exist – dana w polu musi wystąpić w liście haseł wzorcowych, w przeciwnym przypadku ma być wyświetlony komunikat.

Missing – dana w polu nie może wystąpić w liście haseł wzorcowych, w przeciwnym przypadku ma być wyświetlony komunikat.

W obu przypadkach komunikat może przyjąć formę (błędu) ERROR lub WARNING (ostrzeżenia) (patrz: **Forma komunikatu**).

Poniżej przedstawione są przykłady ilustrujące wykorzystanie parametru vmode.

Przykład 20

Załóżmy, że nasz system wyświetli komunikat, jeśli element danych nie istnieje w liście haseł wzorcowych:



```
Vmode = AF; term = exist;
```

Z kolei przyjmijmy, że nasz system wyświetli komunikat, jeśli element danych istnieje w liście haseł wzorcowych:

```
Vmode = AF; term = missing;
```

4. Specyfikacja Poprawności. Parametr ten wyświetla szczegóły określone w valid mode. Dla vmode=AF powinniśmy wyświetlić następujące parametry:

```
Dbn = dbname; pfx = prefix;
```

5. Format komunikatu (w przypadku błędu). Odpowiedzialny jest za utworzenie poprawnego formaty ISIS.

Aby wskazać na odpowiedni typ błędu (ERR), wygenerowany tekst powinien być umieszczony wewnątrz następujących znaczników.

⁸ Opcja ta dostępna jest od wersji 5.0. Pozostałe opcje (INT, REAL, DATE) w chwili pisania podręcznika nie były zaimplementowane.

<ERR> „tekst komunikatu o błędzie”.</ERR>

Jeśli wiadomość ma charakter ostrzeżenia (WARN), tekst powinien być umieszczony między następującymi znacznikami.

<WARN> „tekst komunikatu ostrzegawczego,, </WARN>



UWAGA:

W formacie odpowiedzialnym za komunikat błędu lub ostrzeżenia należy upewnić się, że informacja o każdym komunikacie jest zapisana w osobnym wierszu.

Tak więc zamiast

`<ERR> brak pola autor`

`</ERR>,`

`<ERR> Brak pola tytuł`

`</ERR>,`

format powinien być zapisany następująco:

`<ERR> brak pola autor </ERR>,`

`<ERR> Brak pola tytuł </ERR>,`



UWAGA:

W specyfikacji formatu nie wolno używać znaku #.



Przykład 21

Poniżej jest podany przykład tablicy poprawności pól (składa się z 4 wierszy).



#	
1.	tag= (V30/); select = 0; vmode = AF; term = exist; dbn = CAUTH; pfx = CA=; msgpft = (<ERR>Corporate Author field (page 2): The value V9999 does not exist in the authority list </ERR> /);
2.	tag= V33; select = 0; vmode = AF; term = missing; dbn = FAOBIB; pfx=T=; msgpft = (<ERR> Possibly duplicated record - see Monographic title (page 2): V9999 </ERR>> /);
3.	tag= V72; select = 3; vmode = AF; term = exist; dbn = thes; pfx = E=; msgpft = (<ERR>Main Descriptors field (page 3): The term V9999 does not exist in AGROVOC </ERR> /);
4.	tag= (V75/); select = 0; vmode = AF; term = exist; dbn = thes; pfx = E=; msgpft = (<ERR>Secondary Descriptors field (page 3): The term V9999 does not exist in AGROVOC </ERR> /);

Tablica ta kontroluje zawartość następujących pól:

V30 (corporate author) porównywana jest zawartość danych ze zbiorem wzorcowym, który jest bazą CAUTH (ISIS);

- V33 (monographic title) porównywana jest zawartość danych z roboczą (aktualną) bazą FAOBIB, a w przypadku istnienia danych system generuje komunikat, że właśnie wprowadzany przez nas rekord może już istnieć w naszej bazie;
- V72 (main descriptors) porównywana jest zawartość danych z tezaurem (zagnieżdżone deskryptory);
- V75 (secondary descriptors) porównywana jest zawartość danych z bazą tezaury (każdy deskryptor jest wystąpieniem pola powtarzalnego).

UWAGA:
 Specyfikacja każdego pola powinna być zapisana jako osobny wiersz. W podanych przykładach rozbicie na wiersze jest podyktowane większą przejrzystością zapisu.



5.5.2. Formaty globalnej poprawności

Globalna walidacja posiada strukturę standardowego formatu ISIS, sprawdzane są wszelkie ustalone kryteria ważności rekordu. Jeśli dany warunek nie jest spełniony, format powinien wygenerować jedną z następujących wiadomości:

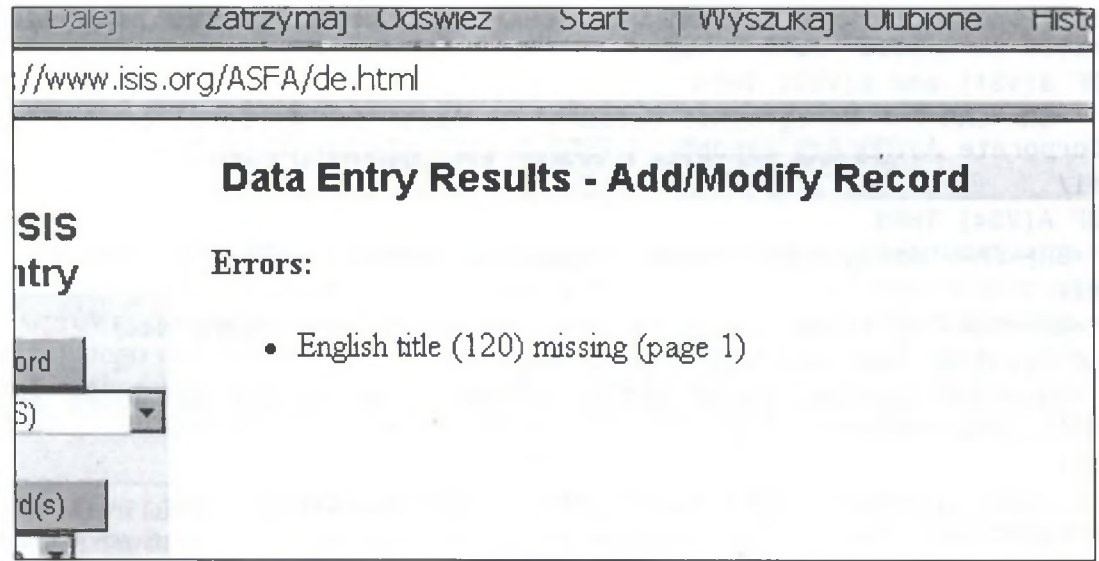
```

<ERR> . . . tekst komunikatu. (o błędzie). . </ERR>
<WARN> . . . tekst komunikatu (ostrzeżenia). . . </WARN>

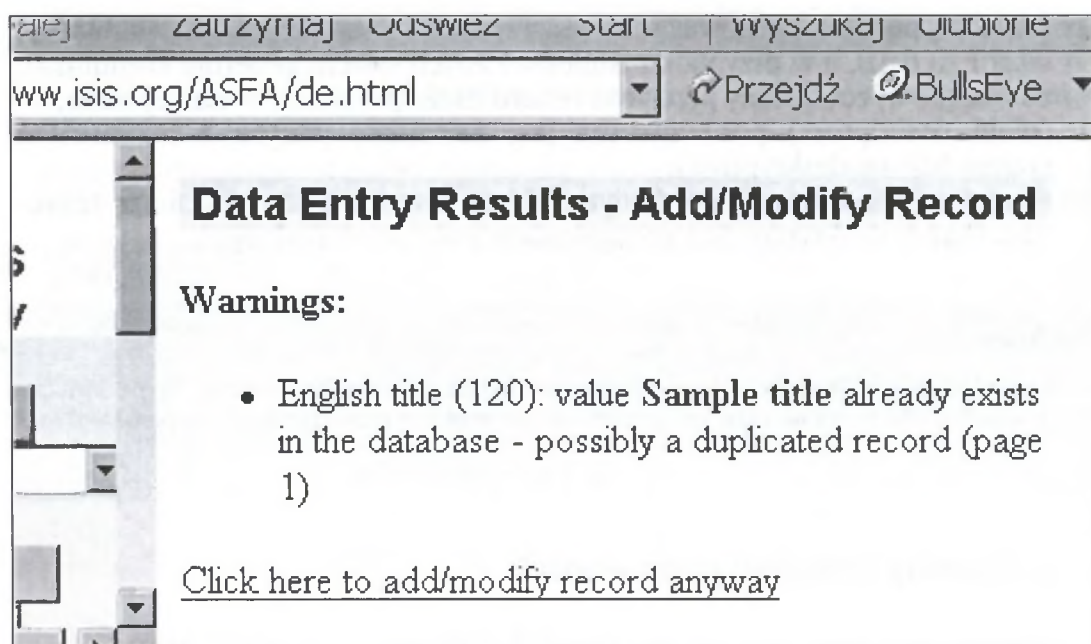
```

W pierwszym przypadku chodzi naturalnie o komunikat błędu (patrz rys. 73), zaś w drugim o ostrzeżenie (rys. 74). Główna różnica między błędem a komunikatem polega na tym, że w przypadku błędu nie jesteśmy w stanie zapisać rekordu, zaś w przypadku ostrzeżenia możemy zapisać rekord do bazy, a komunikat ma wyłącznie charakter ostrzegawczy.

Komunikaty błędu/ostrzeżenia są wyświetlane w polu poniżej tytułu strony (odpowiedzialnej za walidację, rys. 73 i 74). Sposób wyświetlenia komunikatu, kolor i krój pisma, ułożenie tekstu w dużej mierze zależy od projektanta aplikacji.



Rys. 73. Komunikat o błędzie (nie można zapisać rekordu)



Rys. 74. Ostrzeżenie (można zapisać rekord do bazy)

Poniżej przedstawiony jest przykład formatu dla globalnego sprawdzania poprawności :

```
IF (V21:'L') and a(v22) THEN
  '<ERR> <BR>For file assignment '"V21"' -',
  ' the field "Library Call No" is obligatory.</ERR>'
fi/
IF V06='T' and a(v07) then
  '<BR><ERR>"Relator" is "T" - "REL. TRN(07)" field is needed </ERR>'
fi/
IF V11<>'MS' and (p(v41) then
  '<BR><ERR>"SERIAL TITLE" present - "Bibl. Level" must be "MS" </ERR>'
fi/
IF V11='MS' and a(v41) then '<BR><ERR>"Bibl. Level" is "MS" - Serial
title is needed </ERR>' fi/
IF a(v31) and a(v29) THEN
  '<BR><ERR>For Monographic document at least one Personal or
Corporate Author is needed. </ERR>'
fi/
IF A(V34) THEN
  '<BR><ERR>"Monographic title" should be present </ERR>'
fi/
"<BR><ERR>The field 'Lang. of text' is obligatory </ERR>"N44/
if p(v34^e) and (not V44: 'EN') then
  '<BR><ERR>English Title exists - "En" code should be added in
"DOC.LANG"</ERR>'
fi/,
if v54='<(>nd<)' then else "<BR> <ERR>Publication Year must be
present<ERR>"N64
fi/
```

5.5.3. Format kontrolujący usuwanie rekordów

Format sprawdzania usuwanych rekordów posiada typową strukturę formatu ISIS. Zanim system dokona usunięcia rekordu z bazy (wydanego przez polecenie DELETE RECORD), sprawdzane są wszystkie warunki zdefiniowane przez projektanta aplikacji. Jeśli warunki te nie będą spełnione, odpowiedni format spowoduje wygenerowanie komunikatu o błędzie. Poniżej przedstawiamy przykład formatu sprawdzania usuwanych rekordów:

Przykład 22

```
if p(v61) then,  
'<BR><ER> Cannot delete record </ER>'/fi,
```



UWAGA:

Wewnątrz formatów sprawdzających usuwane rekordy możliwe jest użycie wirtualnych pól przechowujących pewne informacje systemowe, jak np. aktualna data, identyfikator użytkownika bazy, adres IP itd. Znaczniki wirtualnych pól są takie same, jak w przypadku formatów dotyczących funkcji on_new oraz on_update czyli:

- 9970** – IP lub adres internetowy zarejestrowanego użytkownika bazy;
- 9960** – Identyfikator użytkownika (używany w momencie rejestracji do systemu);
- 9961** – zdefiniowana w zbiorze Access.def grupa przywilejów (zarejestrowanych) użytkowników (pole to jest powtarzalne – każde wystąpienie pola przechowuje dane o grupie, do której należał użytkownik);
- 9950** – data w formacie Dzień/Miesiąc/Rok.



Dzięki zastosowaniu formatów sprawdzania poprawności oraz zastosowaniu wymienionych wirtualnych pól możliwe jest wprowadzenie ograniczenia dostępu do bazy pewnym użytkownikom. Zaś przy założeniu, że w każdym rekordzie bazy istnieje dodatkowe wirtualne pole informujące o jego właścicielu (owner – użytkownik tworzący dany rekord), odpowiednie formaty sprawdzania poprawności mogą ustalić, czy aktualnie zarejestrowany użytkownik jest uprawniony do dokonywania zmian w danym rekordzie⁹.

5.6. Funkcje ON_NEW i ON_UPDATE

W wielu przypadkach istnieje konieczność tzw. autopoprawek w nowo wprowadzonych lub zmienianych rekordach. Czynności tego typu mogą wiązać się z wprowadzeniem tzw. wirtualnego pola w momencie tworzenia lub uaktualniania danego rekordu. Przykładem takich pól jest identyfikator użytkownika, data zapisania rekordu, przetłumaczenie deskryptorów na inny język, itd. W tym celu zostały zaimplementowane specjalne funkcje ON_NEW oraz ON_UPDATE.

⁹ Wraz z funkcją on_new można zdefiniować pola, które będą automatycznie wypełniane przez system a wprowadzana wartość będzie np. nazwą aktualnie zarejestrowanego i dokonującego zmian w bazie (rekordzie) użytkownika.

Funkcje te mogą być wywołane przez następujące formaty `on_new_pft` oraz `on_update_pft`, do których system odwołuje się w podstawowym zbiorze definicji wprowadzania danych `entry.def`.

Obie funkcje `ON_NEW` oraz `ON_UPDATE` mogą wykonać następujące operacje.

Tabela 15

Operacje wykonywane przez funkcje `ON_NEW` oraz `ON_UPDATE`

ADD	Daje możliwość dodania pola do rekordu.
REPL	Zamienia wartość danego pola na wskazany łańcuch znaków.
DEL	Służy do usunięcia pola z danego rekordu.
SR	Funkcja zamieniająca łańcuch znaków, stosowana w przypadku powtarzalnych pól.

Składnia operacji jest następująca:

```
ADD <tt> tekst </tt>
```

Dodane nowe wystąpienie pola wraz ze znacznikiem `tt`

```
REPL <tt> tekst </tt>
```

Wszystkie wystąpienia pola wraz ze znacznikami `tt` będą usunięte, zaś na ich miejsce będzie wprowadzony wyspecyfikowany tekst.

```
DEL <tt>
```

Usunięte będą wszelkie wystąpienia pól wraz ze znacznikami `tt`

```
SR %%poprzednia wartość %%nowa wartość%%
```

Jak widać nie można zastosować funkcji, jeśli chcemy zamienić każde wystąpienie łańcucha znaków w rekordzie. W szczególności, jeśli pole jest powtarzalne i w każdym wystąpieniu pola chcielibyśmy zastąpić poprzednią wartość nową, to funkcja `REPL` nie będzie działała. W zamian należy zastosować funkcję `SR` (string replace). W przypadku pól powtarzalnych możemy wygenerować sekwencję poleceń, z których każde będzie dotyczyło danego wystąpienia. Na przykład jeżeli format `ON_NEW` wygeneruje tekst:

```
SR %<20>antibiotic</20>%<20>ANTIBIOTIC</20>%%
SR %<20>peniciline</20>%<20>PENICILINE</20>%%
SR %<20>tetracycline</20>%<20>TETRACYCLINE</20>%%
```

będzie to powodowało zamianę zawartości każdego wystąpienia pola o znaczniku `20` z małych liter na duże (operacja będzie wykonana w bazie `ISIS`).

Jeśli parametr `on_new_pft` jest wykorzystywany dla danego typu rekordu i jeśli po sprawdzeniu poprawności wprowadzanych danych nie zostanie wykryty żaden błąd, to po wprowadzeniu rekordu system:

1. Ponownie czyta rekord.
2. Zapisuje datę, identyfikator użytkownika, adres IP w wirtualnych polach.

3. Wykonuje format określony przez parametr on_new_pft; format powinien wygenerować cały szereg poleceń wprowadzających/zmieniających określone pola w rekordzie.

Dzięki określonej sekwencji poleceń system przeprowadza uaktualnianie rekordu. Zmieniony rekord zostaje zapisany.

Format określony przez zmienną on_new_pft jest wywoływany tylko wtedy, kiedy rekord jest nowy. Jeśli istniejący rekord jest aktualizowany, wywoływany jest format on_update_pft i wykonywana jest analogiczna procedura.

Wraz z formatami on_new_pft i on_update_pft możliwe jest wykorzystanie wirtualnych pól, które dostarczają użytecznych danych systemowi. Poniżej dla przypomnienia znaczniki wirtualnych pól.

- 9970 – adres IP użytkownika bazy;
- 9960 – Identyfikator użytkownika (tzw. Login name);
- 9961 – zdefiniowana w zbiorze Access.def grupa przywilejów użytkowników (pole to jest powtarzalne – każde wystąpienie pola przechowuje informacje o danej grupie, do której należał użytkownik);
- 9950 – data w formacie Dzień/Miesiąc/Rok.

Poniżej przykład formatu on_new_pft i on_update_pft.

Przykład 23. Formaty dla funkcji ON NEW, ON UPDATE



'ADD <20> ', MFN(6), '</20> '	Dodanie do pola v20 numeru rekordu
'ADD <20> MON-', MFN(6), '</20> '	Dodanie do rekordu jego identyfikatora w polu v20 (równe MFN). Identyfikator rekordu składa się z ciągu znaków „MON”, po których następuje MFN (Master File Number)
'REPL <20> ', MPU V20, '</20> '	Zamiana wartości w polu z małych na duże litery (po warunkiem, że pole v20 nie jest powtarzalne)
'REPL <20> ', V20, '^a', V21 '</20> '	Dodanie do pola v20 podpola ^a z wartością pola v21 jako zawartością pola v20.
MPU, (SR %<20> , V20, </20>% /)	Jest to powtarzalna grupa, która dla każdego wystąpienia pola v20 uruchamia polecenie SR, zamieniające dany tekst na duże litery.

Dzięki wykorzystaniu funkcji REF w formacie można wykorzystać zakodowane wartości (w postaci kodu) z danego pola oraz dodać do nich rozszerzenia jako wystąpienia innego znacznika (w WWW-ISIS zamiast funkcji REF należy wykorzystać funkcję ISISREF).

Jeśli chcemy przechowywać informację na temat użytkownika, który wprowadził rekord do bazy, to możemy to zrobić wykorzystując następujący format.

Przykład 24. Formaty dla funkcji ON NEW i ON UPDATE (użytkownicy)

'ADD <20> ', V9960, '</20> '	Dodanie identyfikatora użytkownika do pola v20
'ADD <40> ', V9950, '</40> '	Dodanie aktualnej daty do pola v40

Przez dodanie powyższych wierszy do formatu on_update_pft, w podobny sposób możemy kumulować w rekordach „sygnatury” użytkowników aktualizujących rekordy. Jeśli zdecydujemy się, aby umieszczać te „sygnatury” w chronionych polach, w przyszłości będziemy mogli sprawdzać, kto wprowadził, ewentualnie zmienił dany rekord.

5.7. Prawa dostępu, grupy użytkowników

System wprowadzania danych jest wyposażony w mechanizm kontroli dostępu do danych. Możliwe jest zdefiniowanie uprawnień dla poszczególnych grup użytkowników w zakresie zmian, wprowadzania rekordów do baz itd. Wszystkie przywileje są określone przez tablicę praw dostępu do danych przechowywaną w zbiorze `access.def`, umieszczonym w katalogu określonym w zbiorze konfiguracyjnym `isis3w.exe.cfg` w parametrze `EntryAccessFile`.

W zbiorze definiującym tablicę praw dostępu do danych zdefiniowane są wszystkie możliwe grupy przywilejów wraz z użytkownikami oraz informacją, do których grup należą użytkownicy. Każdy użytkownik należy co najmniej do jednej grupy, zaś grupy mają określone prawa dostępu i operacji na danych. Poniżej przedstawiona jest szczegółowo struktura tablicy definicji praw dostępu do danych. Składnia jest podobna do składni tablicy definicji wprowadzania danych i składa się z:

- Sekcji grup użytkowników ISIS.
- Sekcji definiującej grupy użytkowników i prawa dostępu do baz ISIS (arkuszy wprowadzania danych).
- Sekcji definiującej użytkowników.

Sekcja grup użytkowników składa się z wielu sekcji definiujących prawa dostępu. Każdy element definiuje prawa dostępu do arkusza wprowadzania danych. Poniżej podana jest składnia:

```
<isis_group>
  name=nazwaGrupy;
<isis_group_entry>
  worksheet=NazwaArkusza1; rights=r11 r12 r13
</isis_group_entry>
<isis_group_entry>
  worksheet= NazwaArkuszan; rights=r21 r22 r23
</isis_group_entry>
. . . . .
<isis_group_entry>
  worksheet= NazwaArkusza; rights=rn1 rn2 rn3
</isis_group_entry>
</isis_group>
```

W specyfikacji powyżej widoczne są następujące parametry:

Name	– jest nazwą danej grupy
worksheet	– jest nazwą arkusza
rights	– jest parametrem definiującym prawa dostępu. Możliwe są następujące prawa dostępu:
A	– dodaj, utwórz (add new)
M	– modyfikuj (modify)
D	– Usuń (delete)

Składnia sekcji poświęconej użytkownikom wygląda następująco:

```
<isis_user>
  name=użytkownik; password=hasło;
  groups = nazwa1, nazwa2,... nazwak;
</isis_user>
```

Dla każdego użytkownika sekcja **isis_user** określa identyfikator użytkownika, jego hasło oraz listę grup, do których należy.

UWAGA:

Istnieje osobna nazwa użytkownika mająca specjalne znaczenie – **PUBLIC**. Jest ona przeznaczona do definiowania praw użytkowników bez wymogu wprowadzania hasła.



Przykład 25. Definicje praw dostępu

Jeśli na przykład chcielibyśmy umożliwić użytkownikowi PUBLIC składanie np. zamówień (czyli zapisywanie rekordów), zalecane jest zdefiniowanie arkusza ORDER, zaś w tablicy definicji praw dostępu definiujemy następujące elementy:



```
<isis_group>
  name = Guests;
  <isis_group_entry>
    worksheet=ORDER; rights=A;
  </isis_group_entry>
</isis_group>
<isis_user>
  name = public; password=; groups = guests;
</isis_user>
```

Wraz z tą specyfikacją użytkownik może bez rejestrowania się w systemie uzyskać dostęp do arkusza ORDER.

UWAGA:

Dla arkuszy, które zostały udostępnione użytkownikowi PUBLIC standardowa procedura wprowadzania danych działa w sposób ograniczony: rozmiar rekordu ograniczony jest do 2 Kb.



5.8. Inne aspekty interfejsu

5.8.1. Zbiory HTML

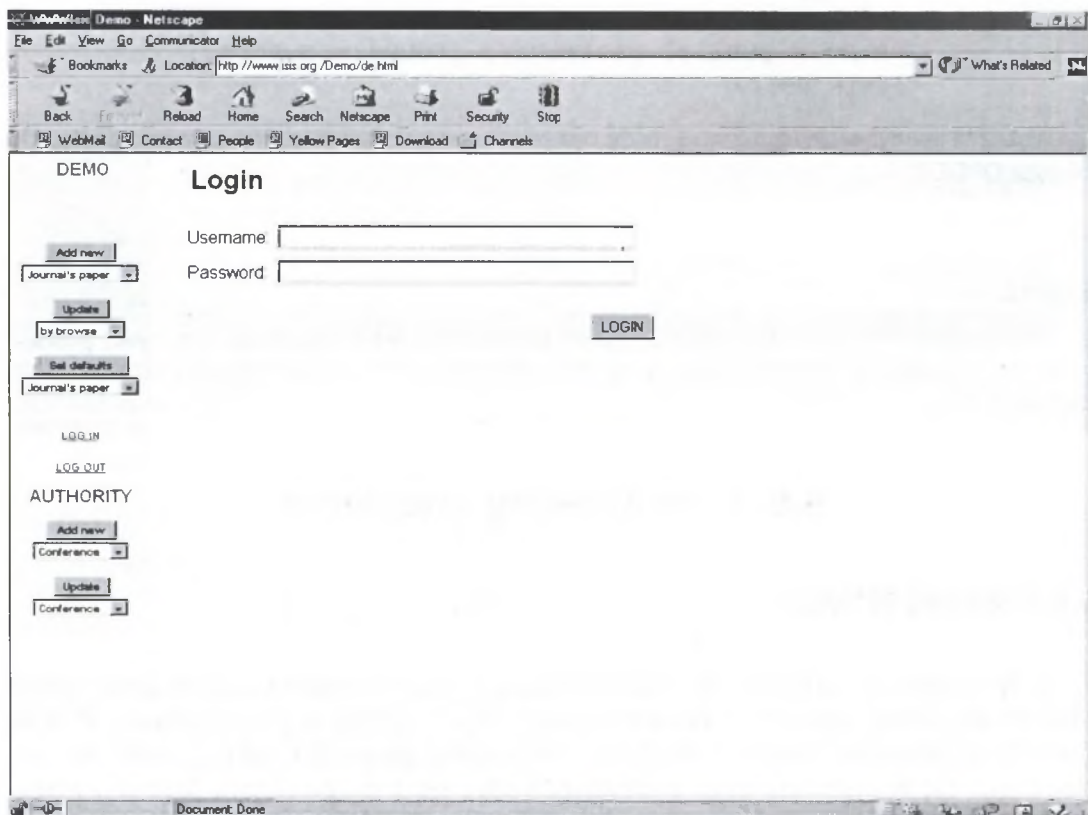
Gdy wszystkie arkusze są zdefiniowane i przetestowane, konieczne mogą okazać się pewne zmiany w zbiorach html. Wyświetlany w przeglądarce WWW formularz powinien zawierać wszystkie niezbędne przyciski, takie jak np. zarejestrowanie się w systemie oraz wyrejestrowanie się z niego (*login*, *logout*), wprowadzanie nowych rekordów, narzędzia do wyboru rekordów w celu ich zmiany itd. Interfejs wyboru rekordów do aktualizacji różni się od interfejsu publicznego systemów bibliotecznych (tzw. OPAC, tj. **O**n-line **P**ublic **A**ccessto **C**atalogue). Na przykład w przypadku projektowania ogólnodostępnego interfejsu w typowej bazie bibliotecznej powinniśmy umożliwić wyszukiwanie zgodne z podstawo-

wymi atrybutami opisu bibliograficznego (autor, tytuł, deskryptory, data konferencji, itd.), zaś interfejs dla pracowników biblioteki powinien zawierać dodatkowo wyszukiwanie rekordów przez MFN (Master File Number), sygnaturę, z tego powodu formaty wyświetlania danych będą się różniły. Ponadto w interfejsie publicznym mogłyby znaleźć się dodatkowe funkcje pozwalające na eksport danych (do formatu ISO2709 lub XML), a w typowym interfejsie wprowadzania danych niezbędne byłyby przyciski wywołujące funkcje: uaktualniania (UPDATE), kopiowania (COPY) oraz usuwania rekordów (DELETE).

Założmy, że interfejs publiczny został już stworzony, w następnym kroku winniśmy dokonać następujących czynności:

- Uaktualnić tablice przeglądu danych WWW-ISIS i utworzyć odpowiadające im formaty wyświetlania.
- Uaktualnić tablice wyszukiwania danych WWW-ISIS i utworzyć odpowiadające im formaty wyświetlania.

W tym podrozdziale zajmiemy się dodatkowymi aspektami interfejsu wprowadzania danych. Najpierw przeanalizujemy strukturę plików HTML systemu wprowadzania danych aplikacji DEMO. Zbiór wywoławczy interfejsu wprowadzania danych to de.html i znajdują się w nim informacje na temat podziału strony wywoławczej na dwie ramki: w jednej wyświetlany jest zbiór tcode.html zaś w drugiej dynamicznie generowana przez isis3w.exe strona rejestracyjna systemu (rys. 75).



Rys. 75. Strona wywoławcza pozwalająca na rejestrację w systemie (login)

Strona pozwalająca na rejestrację w systemie jest wywoływana w lewej ramce strony głównej przez następujące odwołanie hipertekstowe:

```
<FRAME SRC=
"/cgi-bin/isis3w.exe?&search_type=ef_login& page_header=ephmon&
lang=eng&table=mons"
NAME="display">
```

Lewa ramka (patrz rys. 75) przedstawia zestaw przycisków (połączeń) pozwalających na dostęp do podstawowych funkcji systemu. Dzięki nim użytkownik może wybrać m.in. arkusze wprowadzania danych.

5.8.2. Wprowadzanie nowych rekordów

Wybór odpowiedniego arkusza wprowadzania danych, pozwalającego na zróżnicowany opis bibliograficzny (bibliographic worksheet) może być osiągnięty w formie listy wyboru, co w kodzie HTML zapisujemy w następujący sposób:

```
<FORM METHOD=POST ACTION="/cgi-bin/isis3w.exe" TARGET="display"
ENCTYPE="application/x-www/form-urlencoded">
<FONT FACE="Arial,Helvetica" COLOR=green SIZE="2"><B>
<BR>
<FONT FACE="Arial,Helvetica" COLOR=green SIZE="1">
<BR>
  <INPUT TYPE="HIDDEN" NAME="database" VALUE="DEMO">
  <INPUT TYPE="HIDDEN" NAME="search_type" VALUE="ef_new">
  <INPUT TYPE="HIDDEN" NAME="lang" VALUE="eng">
<INPUT TYPE="SUBMIT" VALUE="Add new">
<BR>
<SELECT NAME="de_worksheet" SIZE=1 ">
  <OPTION SELECTED VALUE="ASEE">Journal's paper
  <OPTION VALUE="AMSE">Book's paper
  <OPTION VALUE="MONE">Book
</SELECT>
</FORM>
```

W powyższym kodzie HTML zaznaczony pismem pochyłym parametr odwołujący się do bazy jest opcjonalny. Jeśli jest określony parametr określający domyślną bazę danych w zbiorze entry.def, podanie nazwy bazy jest ignorowane przez system. Wiersze obligatoryjne zaznaczone są pismem wytłuszczonym. Parametr de_worksheet jest definiowany przez końcowego użytkownika dzięki polu typu kombi pojedynczego wyboru.

W podobny sposób możemy zdefiniować arkusz, a w nim odpowiednie pola kombi przez ustalenie wartości domyślnych:

```
<FORM NAME="formdf1t" METHOD=POST ACTION="/cgi-bin/isis3w.exe"
TARGET="display"
ENCTYPE="application/x-www/form-urlencoded">
  <INPUT TYPE="HIDDEN" NAME="database" VALUE="DEMO">
  <INPUT TYPE="HIDDEN" NAME="search_type" VALUE="ef_new">
```

```

<INPUT TYPE="HIDDEN" NAME="de_define_defaults" VALUE="1">11
<INPUT TYPE="HIDDEN" NAME="lang" VALUE="eng">
<INPUT TYPE="SUBMIT" VALUE="Set defaults">
<BR>
<SELECT NAME="de_worksheet" SIZE=1">
  <OPTION SELECTED VALUE="ASEE">Journal's paper
  <OPTION VALUE="AMSE">Book's paper
  <OPTION VALUE="MONE">Book
</SELECT>
</FORM>

```



5.8.3. Uaktualnianie i kopiowanie istniejących rekordów

Przeanalizujmy w formularzu HTML sposób zdefiniowania dostępu do rekordów w celu ich aktualizacji. Mamy tu pole kombi pojedynczego wyboru, z opcjami: przeglądania danych lub ich wyszukiwania.

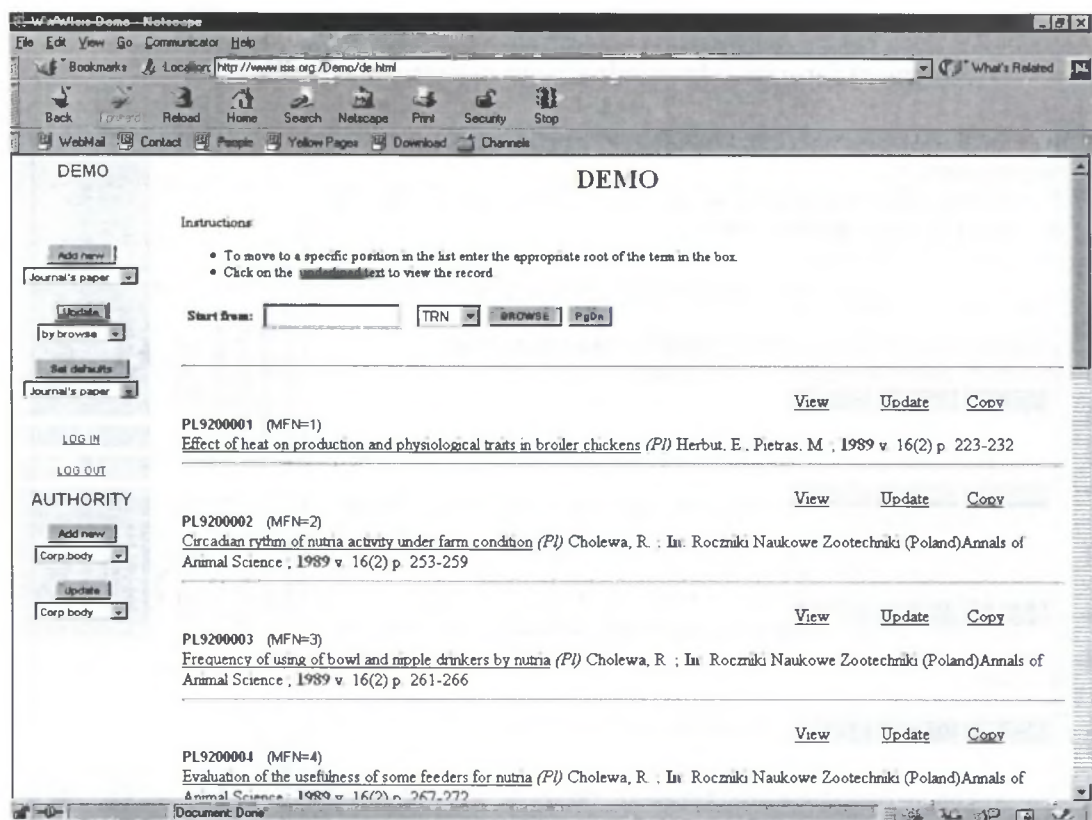
```

<FORM METHOD=POST ACTION="/cgi-bin/isis3w.exe" TARGET="display"
ENCTYPE="application/x-www/form-urlencoded">
  <INPUT TYPE="HIDDEN" NAME="database" VALUE="DEMO">
  <INPUT TYPE="HIDDEN" NAME="table" VALUE="emonm">
  <INPUT TYPE="HIDDEN" NAME="prefix" VALUE="NM:*@ESMONE">
  <INPUT TYPE="HIDDEN" NAME="format_name" VALUE="@ESMONE">
  <INPUT TYPE="HIDDEN" NAME="lang" VALUE="eng">
  <INPUT TYPE="SUBMIT" VALUE="Update">
  <BR>
  <SELECT NAME="search_type" SIZE=1>
    <OPTION SELECTED VALUE="browse">by browse
    <OPTION VALUE="entry">by search
  </SELECT>
</FORM>

```

Jak łatwo zauważyć w przypadku przeszukiwania odwołujemy się do zbioru database.def.eng, a w nim do tablicy emonm. Tablica ta różni się od tablic przeszukiwania danych dostępnych publicznie (mons oraz mona). Także w przypadku przeglądania danych odwołujemy się do tablicy emonm (tym razem w zbiorze browse.def.eng).

Pamiętajmy jednak, że niezależnie od sposobu, w jaki uzyskujemy dostęp do rekordów, format wyświetlania jest zawsze taki sam i jest to @ESMONE (zdefiniowany w oparciu o esmon.pft) . Zanim przeanalizujemy format, przyjrzyjmy się formularzowi uaktualniania danych wyświetlanemu po wybraniu opcji przeglądania danych (BROWSE).



Rys. 76. Strona wyboru rekordów do uaktualnienia

Przykład 26. Omówienie formatu

Strona pokazana na rys. 76 jest generowana dzięki formatowi, którego fragment przedstawia tablica 6. Format ten jest bardzo podobny do formatu @ESMON. We fragmencie przedstawionym w tab. 22 warto zwrócić uwagę na cztery istotne części:



1. Funkcja wyświetlania rekordu w innym formacie – funkcja View (tabl. 6, wiersze 5-4).
2. Funkcja uaktualniania rekordów Update (tabl. 6 wiersze 7-9).
3. Funkcja kopiowania rekordów Copy (tabl. 6 wiersze 11-14).
4. Funkcja kopiowania Copy M->A (tabl. 6 wiersze 15-20).

Przeanalizujemy bardziej szczegółowo funkcję Update (uaktualnianie rekordów). W tym przypadku program cgi (opcja search_type) jest uruchamiany z parametrem „ef_modify”. Łatwo zauważyć, że format wraz z wyrażeniem warunkowym IF określa parametr arkusza w zależności od typu rekordu 12.

Funkcja Copy M->A ma charakter instrukcji warunkowej i działa tylko wtedy, gdy dany rekord jest typu M/MS (czyli jeśli v11.1='M'). Odpowiada on za kopiowanie opisu monograficznego do części monograficznej rekordu „analitycznego” AM (innymi słowy pozwala to przyspieszyć opisanie artykułu z książki, ponieważ opis książki będzie już gotowy). W przypadku kopiowania program cgi jest uruchamiany z parametrem „ef_copy” (opcja search_type). Oczywiście format odpowiedzialny za kopiowanie musi być inny od tego zastosowanego w przypadku funkcji Copy.

Fragment formatu odpowiedzialnego za przeglądanie rekordów

1.	'<TABLE COLS=5 BGCOLOR="e8e8c1" ALIGN="CENTER" WIDTH=98%>'	
2.	'<TR><TD WIDTH=60%><TD WIDTH=10%><TD WIDTH=10%><TD WIDTH=10%><TD WIDTH=10%><TD WIDTH=10%>'	
3.	'<TR><TD><TD><TD>'	
4.	'View'	Połączenie z tym samym rekordem wyświetlanym w innym formacie - View  
6.	'<TD>'	
7.	'<A href="/cgi-bin/isis3w.exe?rec_id=',mf,fn,'&de_worksheet='	
8.	if v11='AS'then'ASEE'else if	
9.	v11.1='A'then'AMSE'else'MONE'fi fi,,	
10.	'&database=demo&search_type=ef_modify&lang=eng">Update'	Wywołanie funkcji aktualizacji rekordu - Update  
11.	'<A href="/cgi-bin/isis3w.exe?rec_id=',mf,fn,'&database=demo&search_type=ef_copy&de_worksheet='	
12.	if v11='AS'then'ASEE'else if	
13.	v11.1='A'then'AMSE'else'MONE'fi fi,	
14.	'&de_copy_init=ECALL&lang=eng">Copy'	Wywołanie funkcji kopiowania rekordu - Copy  
15.	if v11.1='M' then	
16.	'<TD>'	
17.	'<A href="/cgi-bin/isis3w.exe?rec_id=',mf,fn,'&database=demo&search_type=ef_copy	
18.	&de_worksheet=AMSE'	
19.	'&de_copy_init=ECM2A&lang=eng">Copy M->A'	Wywołanie funkcji kopiowania rekordu - Copy M->A  
20.	fi,	
21.	'</TABLE>'	

```
'<FONT FACE="Arial,Helvetica" COLOR=BLUE
SIZE=2><B>'v1'</B>'
```

```
' &nbsp; (MFN='f(mfn,1,0),' )'
```

```
'</FONT><BR>'
```

```
mhl,
```

```
if V11.1='A' then
```

```
'<A href = "/cgi-bin/isis3w.exe?rec_id='
```

```
mfn,
```

```
'&database=demo&search_type=link&table=mona
&lang=eng&format_name=EFMONE">'
```

```
v120,if a(v120) then v122,if a(v122) then v124,if
a(v124) then v126 fi,fi,fi,'</A> '/
```

```
"<I>("v180")</I> "/
```

```
(l, |+v100^n)/
```

```

if p(v220) or p(v420) then '<B>; In: </B>',
v220,
if p(v220) and p(v420) then '; ' fi,
v420/
fi,
if p(v130) then '<I>',
if p(v100^n) or p(v220) or p(v420) then '; ' fi,
v130, ", "v131, ", "v133
'</I>',
fi/
if (p(v100^n) or p(v220) or p(v420) or p(v130))
and (p(v110) or p(v262) or p(v261) or p(v263) or
p(v463)) then '; ' fi/
(v110|; |)/
v262", ",v261", ", "<B>"v263"</B>"/
if a(v263) then "<B>"v463"</B>"fi/
(v470)/

Else
'<A href = "/cgi-
bin/isis3w.exe?rec_id=',mfn,'&database=demo
&search_type=link&table=mona&
lang=eng&format_name=EFMONE">',
v220,if a(v220) then v222,if a(v222) then v224,if
a(v224) then v226 fi,fi,fi,'</A> '/
"<I>("v280")</I> "/
(|, |+v200^n)/
if p(v420) then '<B>; In: </B>',
v420/
fi,
if p(v230) then '<I>',
if p(v200^n) or p(v420) then '; ' fi,
v230", "v231, ", "v233
'</I>',
fi/
if (p(v200^n) or p(v420) or p(v230))
and (p(v210) or p(v262) or p(v261) or p(v263))
then '; ' fi/
(v210|; |)/
v262", ",v261", ", "<B>"v263"</B>"/
fi,
'<HR>'

```

Jak widać, niektóre z funkcji wywoływane są w formacie jako instrukcje warunkowe, w zależności od typu rekordu. W pewnych sytuacjach może okazać się konieczne bardziej precyzyjne ograniczenie dostępu do niektórych funkcji systemu.

Jeśli na przykład w rekordzie przechowujemy identyfikator użytkownika („twórcy rekordu 13”), to możemy wyświetlić funkcję Update warunkowo – będzie ona wyświetlana jedynie wówczas, kiedy bazę przegląda użytkownik o odpowiednim identyfikatorze. Rozwiązanie takie jest możliwe dzięki tzw. *wirtualnym polom*.

9970 – adres IP użytkownika bazy;

9960 – identyfikator użytkownika (tzw. Login name);

9961 – zdefiniowana w zbiorze Access.def grupa przywilejów (zarejestrowanych) użytkowników (pole to jest powtarzalne – każde wystąpienie pola przechowuje jedną dane o danej grupie, do której należał użytkownik);

9950 – data w formacie Dzień/Miesiąc/Rok;

Tak więc, jeśli na przykład pole v20 jest używane do przechowywania informacji o „twórcy rekordu”, to możemy uzyskać kontrolę właściciela rekordu przez porównanie pola v20 z polem v9960:

```
If v9960=V20 then
'<A href="/cgi-bin/isis3w.exe?rec_id=',mfn,'&de_worksheet=',
  if v11='AS'then'ASEE'else if v11.1='A'then'AMSE'else'MONE'fi fi,,
  '&database=demo&search_type=ef_modify&lang=eng">Update</A>',
fi
```



UWAGA 1:

Zaleca się, aby pole twórcy rekordu było tworzone automatycznie przez system i by było chronione przed zapisem z zewnątrz. Oznacza to, że pole takie powinno być automatycznie generowane przez funkcję ON_NEW. Jeśli warunek ten nie będzie spełniony, to istnieje obawa, że wprowadzona manualnie wartość nie będzie zgadzała się z rzeczywiście zarejestrowanym podczas sesji użytkownikiem systemu.



UWAGA 2:

Jeśli w systemie nie zostanie zarejestrowany żaden użytkownik, to wartość w polu 9960 będzie ustawiona na unknown (nieznany).

Kontrola IP

Jeśli funkcja jest planowana dla określonej grupy użytkowników, którzy mogą być identyfikowani przez numer IP w sieci, to można zaprojektować określony przycisk, który warunkowo będzie wybierał określony arkusz dla danego numeru IP.

Kontrola Daty

Jeśli dana funkcja ma być wykonana tylko w określonym czasie, to możliwe jest wykorzystanie pola v9950 (a także wybranych dat). Na przykład operacja kontroli daty może być wykorzystana w celu umożliwienia dostępu do bazy na określony czas (próbny okres) wybranym użytkownikom.

5.8.4. Formaty zatwierdzania danych oraz uaktualnianie rekordów

Jedną z pożytecznych cech systemu jest możliwość wyświetlania zawartości rekordów zaraz po zapamiętaniu w systemie. Rekordy mogą być wyświetlone, jeśli definicja arkusza wprowadzania danych w zbiorze entry.def zawiera parametr proof_pft. Inną użyteczną cechą systemu jest fakt, że po wyświetleniu zmienionego rekordu możliwy jest powrót do listy ostatnio uaktualnianych rekordów w celu pobrania następnego rekordu do uaktualnienia.

Prześledźmy preferowaną kolejność działań dotyczącą uaktualniania jednego lub więcej rekordów:

1. Najpierw dokonujemy przeszukiwania lub przeglądania danych w celu odnalezienia rekordu(ów) w celu ich uaktualnienia.
2. Aktualizujemy rekord z wyświetlonej listy.
3. Sprawdzamy rekord wyświetlony w formacie zatwierdzania danych.
4. Powracamy do listy wyświetlanych rekordów (czyli do ostatniej strony, z której pobraliśmy rekord do uaktualnienia) w celu kontynuacji procesu uaktualniania rekordów.

Zwykle nie jest łatwo zaprogramować odpowiedni scenariusz w architekturze systemów klient-serwer, w szczególności w technologii cgi, gdyż programy cgi nie „pamiętają” przebiegu danej sesji.

UWAGA:

Aby można było zaimplementować możliwość powrotu do poprzedniej sesji wprowadzony jest specjalny parametr stanu ('state'), który umożliwia systemowi WWW-ISIS powrót do poprzedniego stanu. W naszym przypadku chodzi o pole o znaczniku 9940. Konieczne jest dodanie parametru de_back do formatu przeglądania/przeszukiwania

```
( 'de_back=' ,v9940) ;
```

Konieczne jest zdefiniowanie formatu umożliwiającego warunkowy powrót do listy (powrót daje nam możliwość odwołania się do pola v9940).

Przykład 27. Zatwierdzanie i uaktualnianie danych

Poniżej podajemy przykład formatu wyświetlania pozwalający na powrót do poprzedniego stanu (lista rekordów do uaktualnienia) fragment z tablicy 6.

```
'<A href="/cgi-bin/isis3w.exe?rec_id=',mfn,'&de_worksheet=',  
if v11='AS'then'ASEE'else if v11.1='A'then'AMSE'else'MONE'fi fi,,  
'&database=demo&search_type=ef_modify&lang=eng&de_back=' ,v9940">  
Update</A>',
```

Część odpowiedzialna za przechowywanie informacji o stanie sesji zaznaczona jest pismem wytłuszczonym (..&de_back=',v9940..).



Format zatwierdzania danych

Fragment formatu odpowiedzialnego za zatwierdzanie danych (przypisanych do klawiszy funkcyjnych) jest pokazany w tab. 16.

Tabela 16

Fragment formatu zatwierdzania danych

1.	'<HR>'
2.	'<TABLE COLS=5 BGCOLOR="e8e8c1" ALIGN="CENTER" WIDTH=98%>',
3.	'<TR><TD WIDTH=60%><TD WIDTH=10%><TD WIDTH=10%><TD WIDTH=10%><TD WIDTH=10%>',
4.	'<TR><TD><TD>',
5.	if p(v24) and p(v9940) then
6.	'<A href="/cgi-bin/isis3w.exe?search_type=ef_back&de_back=',v9940,
7.	'">Return to selection list',
8.	fi,'<TD>',
9.	'<A href="/cgi-bin/isis3w.exe?rec_id=',mf,'&de_worksheet=',v4"E",
10.	'&database=asfa&search_type=ef_modify&lang=eng','">Update',
11.	'<TD>',
12.	'<A href="/cgi-bin/isis3w.exe?rec_id=',mf,'&de_worksheet=FULL',
13.	'&database=asfa&search_type=ef_modify&lang=eng','">Update (complete)',
14.	'<TD>',
15.	'<A href="/cgi-bin/isis3w.exe?rec_id=',mf,
16.	'&database=asfa&search_type=ef_copy&de_worksheet=',v4"E",
17.	'&de_copy_init=ECALL&lang=eng">Copy',
18.	'</TABLE>',

Część formatu odpowiedzialna za zatwierdzanie danych wyświetlona jest w tab. 16, wiersz 5-8. Zauważmy, że funkcja powrotu „Return to Selection List” (wiersz 7) jest stosowana warunkowo, tzn. tylko wtedy, gdy pola v9940 i v24 są obecne w rekordzie. W tym przykładzie zakładamy, że pole v24 jest datą ostatniego uaktualnienia rekordu („Date of last update”) i jest dostępne jedynie wtedy, gdy rekord był zapisywany jako uaktualniany (dzięki formatowi ON Update, który tworzy pole v24, podczas gdy format ON New nie robi tego – daje to gwarancję, że format ten nie wyświetli połączenia powrotu w przypadku, gdy dokończymy wprowadzać nowy rekord, natomiast wyświetli tę funkcję, gdy kończymy aktualizować dany rekord).

UWAGA:

Warunek określony w formacie w wierszach 5-8 jest obligatoryjny, w przeciwnym bowiem razie funkcja "Return to selection list" byłaby dostępna również przy wyświetlaniu nowo wprowadzonych rekordów do bazy.



5.8.5. Rejestrowanie i wyrejestrowywanie użytkowników

Na koniec nasz interfejs powinien zawierać funkcję rejestrowania użytkownika w systemie oraz dawać możliwość wyrejestrowywania się z systemu.

URL:

Aby połączenie w adresie przeglądarki pozwalało na rejestrację w systemie (login), konieczny jest następujący zapis:

```
search_type=ef_login
```



URL:

Aby połączenie w adresie przeglądarki pozwalało na wyrejestrowanie się z systemu (logout) konieczny jest następujący zapis:

```
search_type=ef_logout
```



Przykład 28. Odwoływanie się do funkcji rejestrowania się (login)

```
<A HREF="/cgibin/isis3w.exe?database=DEMO&search_type=ef_login& page_header=ephmon&lang=eng&table=mona" TARGET="display">
```



Przykład 29. Odwoływanie się do funkcji wyrejestrowania się (logout)

```
<A HREF="/cgibin/isis3w.exe?database=DEMO&search_type=ef_logout&page_header=ephmon&lang=eng&table=mona" TARGET="display">
```



5.9. Poczta elektroniczna, strony kodowe

5.9.1. Narzędzia do poczty elektronicznej

System WWW-ISIS daje również możliwość generowania poczty elektronicznej w momencie zapamiętywania rekordu (nowego lub aktualizowanego). Może to być wykorzystane w wielu przypadkach organizowania współpracy różnych grup użytkowników (tzw. Workflow systems). Przykładowe możliwe zastosowania to:

1. Poinformowanie administratora bazy o wprowadzaniu nowych rekordów bazy. Przykładem takiej aplikacji byłaby baza na temat źródeł informacji (instytucji), w której zainteresowane firmy, czy instytucje informacyjne same mogą zarejestrować swoje usługi, zaś informacja o tym fakcie powinna być automatycznie dostarczona do centrum koordynacyjnego.

2. Informowanie administratora tezaursu o nowych propozycjach deskryptorów; propozycje mogą być zgłaszane przez wybranych użytkowników, a administrator powinien je śledzić i ewentualnie wprowadzać w życie.

3. Implementacja funkcji zamówień – odpowiednio zdefiniowany formularz pozwoliłby końcowemu użytkownikowi wysyłać żądania przysłania konkretnego dokumentu; zapisanie żądania w bazie danych w formie ustrukturalizowanej pozwoli tworzyć statystykę zapotrzebowania na dokumenty.

4. Zintegrowany system wypożyczeń, w którym rejestracja zwrotu książki, może automatycznie generować e-mail do oczekującego w kolejce użytkownika z informacją, że książka jest już dostępna.

Przykładów takich można oczywiście mnożyć znacznie więcej. Z punktu widzenia końcowego użytkownika funkcje tego typu są implementowane w postaci hipertekstowych połączeń. Istotną kwestią jest stworzenie odpowiedniego połączenia, którego wywołanie tworzy nowy rekord w bazie ISIS. W tym momencie rekord jest zapisywany w bazie, a program cgi wysyła wiadomość pocztą elektroniczną (np.: do pracowników biblioteki, administratora bazy, centrum informacyjnego itd.), gdzie odbywa się dalsze przetwarzanie danych związanych z rejestracją rekordu.

Wraz z funkcją obsługi poczty **mail_fmt** możliwe jest żądanie kopiowania (**Copy**), wprowadzania nowych rekordów (**New**) oraz ich uaktualniania (**Update**). Jeśli ten parametr został odpowiednio określony, to wraz z pomyślnym zapisem rekordu (**Copy**, **New** lub **Update**) WWW-ISIS wykorzystuje określony format w celu utworzenia wiadomości w poczcie elektronicznej.

Z punktu widzenia użytkownika komunikacja z systemem jest taka sama, jak w przypadku innych funkcji aktualizacji bazy danych. W momencie wybrania funkcji (kliknięcia na połączeniu) system wykonuje następujące operacje:

1. Tworzony jest formularz (pusty lub z predefiniowanymi wartościami).

2. Jeśli formularz został prawidłowo wypełniony i wysłany do serwera, system dokonuje sprawdzenia poprawności wprowadzonych danych.

3. Jeśli nie zostały wykryte żadne błędy, rekord jest zapisany, następnie aktywowany jest format **mail_format**. Format generuje typowy list w poczcie elektronicznej wraz z koniecznymi parametrami, takimi jak: **from**, **to**, **reply-to**, itd., zaś serwer poczty elektronicznej wysyła dalej tak sformatowany list.

Założmy, że chcemy zaimplementować system zamówień książek. Zamówienia chcemy przechowywać w specjalnej bazie danych. Na poziomie implementacji takiej funkcji konieczne są następujące kroki:

KROK 1

CDS/ISIS
FST FDT

WWW-ISIS



Stworzenie dodatkowej bazy ISIS w której przechowywany byłby chronologiczny zapis zamówień ("request log").

KROK 2

CDS/ISIS
FORMATY

WWW-ISIS



Zdefiniowanie formularzy wprowadzania danych oraz wszystkich koniecznych narzędzi sprawdzania poprawności wprowadzonych danych.

KROK 3

<HTML>

WWW-ISIS



Dodanie połączeń hipertekstowych do głównych formularzy wyświetlania danych.

Przykład 30. Funkcja mail_fmt w bazie zamówień

Przy założeniu, że baza ISIS jest określona jako ORDERS (zamówienia), zaś arkusz wprowadzania danych nazywa się ORDER, a email.pft jest formularzem dla obsługi poczty elektronicznej, to powinniśmy stworzyć następujące połączenie hipertekstowe:



```
<A href=
"/cgi-bin/isis3w.exe?search_type=ef_new&de_worksheet=order&
database=orders& de_mail_pft =email&lang=eng">
Order the material </A>
```

Zaleca się, aby kopiować pewne dane z rekordu źródłowego do rekordu w bazie ORDER (takie jak sygnatura, autor, tytuł itd.). Dlatego lepiej jest użyć funkcji copy.

```
<A href="/cgi-bin/faobib.exe?search_type=ef_copy&de_copy_init=
ECLIB&de_worksheet=order&database=faobib&rec_id=000839&de_mail_pft
=email&lang=eng"> Order the material </A>
```

Wskazany format będzie generował listy w poczcie elektronicznej o następującej strukturze:

mail server IP (nie nazwa, ale IP)

mail server port (zwykle 25)

To: <adres> [powinien być w wierszu 3, wielkość liter w polu "To:" jest ważna]

From: <adres> [powinien być w wierszu 4, wielkość liter w polu, "From:" jest ważna]

Reply-to: <adres>

Subject: <temat listu>

<nowy wiersz>

<nowy wiersz>

część zasadnicza listu

<pusty wiersz>

Poniżej pokazano, jak może wyglądać przykładowy list w poczcie elektronicznej wygenerowany przez WWW-ISIS.

#	
1.	148.81.64.3
2.	25
3.	To: library@fao.org
4.	From: brown@minagr.gov.uk
5.	Reply-to: brown@minagr.gov.uk
6.	Subject: request for a book
7.	
8.	Dear Sir/Madam
9.	Please send me the document entitled:
10.	Summary report of and papers presented at the tenth session of the Working Party on Fish Technology and Marketing. Asia-Pacific Fishery Commission. Working Party on Fish Technology and Marketing. Sess. 10, by James, D.G., (ed.).
11.	Accession Number: 379509.
12.	The document should be sent under address:
13.	Abacki Piotr
14.	Elektronika
15.	Teatralny 234
16.	Poland
17.	
18.	Sincerely Yours
19.	Abacki Piotr

W powyższym liście zostały użyte następujące parametry:

148.81.64.3 – numer IP komputera, na którym ulokowany jest serwer poczty i do którego musi odwoływać się program WWW-ISIS. Zaleca się, żeby był to najbliższy położony serwer poczty elektronicznej (względnie taki, do którego dostęp jest najszybszy i najłatwiejszy).

25 – numer portu serwera poczty, do którego odwołuje się WWW-ISIS, zwykle jest to port 25.

To: library@fao.org

From: brown@minagr.gov.uk

Reply-to: brown@minagr.gov.uk

Subject: request for a book



UWAGA:

Dla poprawnego funkcjonowania poczty w systemie WWW-ISIS konieczne jest dodanie odpowiedniego wiersza do zbioru isis3w.exe.cfg definiując katalog buforowy dla poczty WWW-ISIS:

```
OrdersMailSpool np:C:\waicent1\spool
```

Podanie katalogu buforowego jest konieczne. W przypadku, gdy wystąpią trudności z połączeniem z serwerem poczty, system WWW-ISIS będzie jednak tworzył przesyłki pocztowe i składował je w katalogu buforowym. W przypadku nawiązania łączności z serwerem przesyłki mogą być przesłane do serwera automatycznie.

Dzięki istnieniu specjalnego wirtualnego pola (V9970) przechowującego adres IP użytkownika, możliwe jest dodatkowo żądanie sprawdzania uprawnień do korzystania z funkcji obsługi poczty elektronicznej.

5.9.2. Strony kodowe

WWW-ISIS pozwala utrzymywać bazę wraz z obcymi znakami diakrytycznymi. Aby zachować znaki diakrytyczne w zbiorze odwróconym należy zdefiniować następujące tablice w ISIS: ISISAC oraz ISISUC. Podobnie jak w Micro CDS/ISIS, tablica ISIAC powinna zawierać listę wszystkich znaków, które muszą być wzięte pod uwagę w momencie budowania zbioru odwróconego, zaś tablica ISIUC jest wykorzystywana w celu zamiany małych liter na duże (jedynie pierwsza część tablicy, czyli pierwszych 255 znaków jest brane pod uwagę).

Tablice powinny odpowiadać określonej systemowi kodowania: czyli może być to albo standard znaków w Windows, albo w ISO. W celu publikacji bazy w Internecie zaleca się stosowanie standardu ISO. Tablice ISIS powinny być przechowywane w katalogu baz danych ISIS. Jeśli nie została zdefiniowana żadna tablica stosowane są domyślnie tablice: ISISAC oraz ISISUC (zgodnie z rekomendacjami UNESCO). Jeśli została zdefiniowana tablica ISISAC, to w konsekwencji konieczne jest zdefiniowanie tablicy ISISUC w celu właściwego funkcjonowania opcji przeglądania danych oraz ich wyszukiwania.

5.10. Zbiory HTP i konfiguracyjne

Wszystkie formularze są generowane automatycznie przez specjalne zbiory HTP. Podobnie arkusze wprowadzania danych są budowane za pomocą zbiorów HTP. Zbiory HTP pełnią rolę analogiczną jak w systemie wyszukiwania danych (por. rozdz. 4. Wyszukiwanie w WWW-ISIS) i dostarczają wzorca, wedle którego zbiory HTML będą tworzone automatycznie.

Zbiory HTP w WWW-ISIS

search.htp	wzorce wykorzystywane w celu budowania formularzy wyszukiwawczych;
result.htp	wzorce do tworzenia stron z wynikami wyszukiwań;
thes.htp	wzorec do tworzenia strony z polem do wprowadzania deskryptorów wyszukiwawczych z tezaurusa;
view.htp.	wzorec do tworzenia strony historii wyszukiwań oraz strony do wyszukiwania w języku wyszukiwawczym ISIS;
browse.htp	wzorce wykorzystywane w celu budowania formularzy przeglądania danych;
error.htp	wzorce wykorzystywane w celu przechowywania wszystkich komunikatów systemu;
e_forms.htp	wzorce wykorzystywane w celu budowania formularzy arkuszy wprowadzania danych;
e_browse.htp	wzorce wykorzystywane w celu budowania formularzy arkuszy wyświetlania danych do uaktualniania.

Pięć pierwszych plików zostało już szczegółowo omówionych w rozdziale 4. Poniżej skoncentrujemy się jedynie na szablonach e-forms.htp oraz e-browse.htp

W standardowej instalacji nie ma potrzeby modyfikowania zbiorów HTP. Jedyny fragment kodu, jaki winien być zmieniony w zbiorach e-forms.htp oraz e-browse.htp to wiersz 3:

```
#3 <TITLE>DEMO Data Entry</TITLE>
```

Ponadto w szczególnych przypadkach pewne modyfikacje mogą być dokonane odnośnie kolorów, proporcji tabel, dodatkowego tekstu wyjaśniającego itd. Zbiór konfiguracyjny wraz ze wszystkimi parametrami został omówiony w Aneksie A.

5.11. Ćwiczenia



Ćwiczenie 22

Znajdź w katalogu macierzystym zainstalowanej aplikacji zbiór entry.def, następnie otwórz go za pomocą edytora tekstowego (np. Notepad lub WordPad) oraz znajdź arkusz ASEE. Uruchom system wprowadzania danych DEMO i postaraj się wprowadzić nowy rekord jako „Journal from paper” (artykuł z czasopisma), porównaj wyświetloną na ekranie nazwę arkusza z nazwą określoną w tablicy definicji wprowadzania danych (parametr title).

Zmień parametr odpowiedzialny za nazwę (title) w zbiorze entry.def, zapisz zbiór, przeładuj arkusz i sprawdź zmiany na ekranie.

WWW-ISIS

Ćwiczenie 23

Przygotuj zbiór HTML, w którym będą widoczne osobne przyciski dla rejestrowania użytkowników i ich wyrejestrowywania.

<HTML>

WWW-ISIS

Ćwiczenie 24

Otwórz system wprowadzania danych korzystając z przeglądarki Netscape (może być Internet Explorer), wybierz nowy rekord (new record) „journal from paper”, na stronie 5 będzie znajdowało się pole abstraktów (Abstract). Następnie w specyfikacji arkusza ASEE w aplikacji DEMO w zbiorze entry.def znajdź opis tego pola i dodaj parametr repeat=yes. Zapisz zmiany i przeładuj stronę w przeglądarce. Sprawdź, jak funkcjonuje przycisk ADD NEW.

WWW-ISIS

Ćwiczenie 25

Sprawdź formaty ON_NEW oraz ON_UPDATE dla arkuszy „Paper from Journal” oraz wytłumacz ich działanie (w systemie DEMO).

CDS/ISIS
FORMATY

Ćwiczenie 26

Przeanalizuj format EBTK z bazy THES (aplikacja DEMO):

CDS/ISIS
FORMATY

<HTML>

```
'<TR><TD VALIGN=TOP><INPUT TYPE="CHECKBOX" NAME="de_browse_name"
OnClick="cb_click(this);" VALUE="'
if p(v001) then v001 else V003+|;| fi,
'"><TD>'
if p(V1) then,
V9980,
': &nbsp; ',
'<A HREF = "/cgi-bin/isis3w.exe?rec_id=',mf,
'&database=THES&format_name=EBR2_0&search_type=link&table=dsc2&lan
g=eng">',
V001,'</A>'/,
Else if p(V008) then,
'<FONT COLOR="#ff0000">',<I> ',V008,</I> </FONT><BR>',
'&nbsp; &nbsp; &nbsp; <B> use: &nbsp; </B>',
(if p(V003)then
|<B>&nbsp; and &nbsp; </B>|+d003,
'<A HREF="/cgi-bin/isis3w.exe?query=E:',V003,
'&database=THES&search_type=query&table=dsc2&lang=eng&format_name=
EBR2_0">',
V003,mpl,'</A>&nbsp; &nbsp; ',
fi)/,
fi,fi,
' <BR>',
```


Aneksy

Aneks A

Konfiguracja – isis3w.exe.cfg

Możemy założyć, że czytelnik miał już kontakt ze zbiorem isis3w.exe.cfg, przeprowadzając pierwszy test swojego systemu, choć wówczas nie wnikalismy w znaczenie wszystkich parametrów. Teraz przyjrzymy się wszystkim tym parametrom bardziej szczegółowo. Niech przykładowym zbiorem konfiguracyjnym będzie zbiór z aplikacji DEMO.

Zbiór konfiguracyjny jest przechowywany w katalogu skryptów cgi (jest to przeważnie katalog cgi-bin) wraz z programem cgi (skrypcem WWW-ISIS). Zasada nazewnictwa zbiorów konfiguracyjnych jest następująca: do nazwy samego programu dodaje się dodatkowe rozszerzenia „.cfg”, czyli dla programu isis3w.exe mamy isis3w.exe.cfg. Konfiguracje niektórych serwerów wymagają, aby program cgi był zawsze wywoływany z rozszerzeniem „.exe”. Aby móc korzystać z kilku różnych niezależnych aplikacji WWW-ISIS na jednym komputerze można albo zdefiniować kilka wirtualnych serwerów (programy cgi będą rozmieszczone w różnych katalogach), albo umieścić programy cgi w jednym katalogu cgi, ale pod różnymi nazwami zmieniając przy tym nazwy programów konfiguracyjnych i odpowiednie odwołania w stosownych zbiorach html.

Pamiętajmy, że w przypadku wersji „demo” zbiór isis3w.exe.cfg jest tworzony automatycznie w czasie instalacji i jego ustawienia zależą w dużej mierze od ustawień serwera.

Tabela 18

Parametry przykładowego zbioru konfiguracyjnego isis3w.exe.cfg¹

Parametry i przykładowe wartości w zbiorze isis3w.exe.cfg		Ważność parametru
# ----- # ISIS3W demonstration config file # -----		5
DebugLevel	0	4
DebugLogFile	D:\temp\isis3w.log	
DB_name	demo	3
RestrictedTermMinLength	2	
CookieDomain	faecd.fao.org	2
CookiePath	/cgi-bin/	2
QueryMaxRows	600	4
HttpFilePath	c:\Xitami\Web\demo\http	1
IncludeFilePath	c:\Xitami\Web\demo\http\eng	1
SearchDefinitionFile	c:\Xitami\Web\demo\database.def	1
EntryDefinitionFile	c:\Xitami\Web\demo\entry.def	6

¹ Należy pamiętać, że wartości w rzeczywistym zbiorze konfiguracyjnym, z którym przyjdzie pracować czytelnikowi, mogą się w sposób znaczny różnić od zamieszczonego tutaj przykładu.

BrowseDefinitionFile	c:\Xitami\Web\demo\browse.def	1
HighlightSearchTerms	1	4
LatinConvOn	0	
LatinConvTerms	1	
DefaultFormat	@ESMON	1
IsisForceType	0	3
SearchPaging	1	
BrowsePaging	1	
PagingMaxRows	10	
PagingMaxPages	10	
# Enable multi database search		5
IsisMultiDb	0	3
IsisMultiDbUser	0	
IsisMultiDbFile	D:\isis3w\WEB\demo\multidb.def	
# Below is the name of the directory containing ISIS database files.		5
IsisDataDir	D:\faocd\demc\	1
# ----- # Isis documents location URL prefix for single DB mode # For multi DB mode this URL prefix is set for each database in # multi DB definition file.		5
IsisDocLocation	file:///E:/data/docs/	6
# Number of occurrences of isis document location field		5
IsisDocLocationOcc	0	6
IsisTempFileLife	100	1
IsisWorkDir	C:\Windows\Temp\	
AndOr	1	6
FieldAndOr	1	
Sort	1	
Format	1	
TermSeparator	;	
ListCheckBox	1	

Nie wszystkie parametry zbioru konfiguracyjnego są równie istotne. Aby uświadomić czytelnikowi różny poziom ważności danego parametru w tab. 18, zastosowaliśmy w kolumnie obok krótką informację na temat ważności danego parametru:

- 1 – Parametr obligatoryjny bardzo ważny, ustawiany przez projektanta interfejsu WWW-ISIS.
- 2 – Parametr, którego ustawienie może być niezbędne w przypadku korzystania z historii wyszukiwań oraz tezaury.
- 3 – Parametr, który może być ustalany przez projektanta, ale pozostawienie wartości domyślnych (tworzonych w czasie instalacji) nie powinno spowodować zaburzeń pracy systemu.

- 4 – Oznacza parametry, które administrator powinien zapamiętać na wypadek problemów w pracy systemu.
- 5 – Komentarz.
- 6 – Parametry nieistotne.

W tab. 19 czytelnik znajdzie szczegółowe omówienie każdego z parametrów zbioru konfiguracyjnego.

Tabela 19

Omówienie wszystkich parametrów zbioru konfiguracyjnego (isis3w.exe.cfg)

Parametr	Opis	Ważność parametru
AndOr	Dla AndOr równego 1 wyświetlone zostanie pole typu lista (pozwalające na łączenie terminów operatorami Boolowskimi: AND, OR); dla AndOr równego 0 nie zostanie wyświetlone pole a terminy wyszukiawcze będą domyślnie łączone w zapytaniu operatorem OR.	6
BrowseDefinitionFile	Jest to katalog, w którym przechowywany jest zbiór browse.def, służący do definiowania tablicy przeglądania danych.	1
BrowsePaging	Gdy parametr ten wynosi 1, rezultaty przeglądania będą podzielone na strony, zaś rozmiar strony jest zdefiniowany przez PagingMaxRows.	3
CookieDomain	Paramatr ten jest istotny dla poprawnego funkcjonowania historii wyszukiwań oraz tezaurya i określa nazwę domenową danego komputera. Należy pamiętać o właściwych przeglądarkach pozwalających na korzystanie ze zmiennych typu „cookie”.	2
CookiePath	Patrz wyżej (przeważnie jest to katalog /cgi-bin/)	2
DB_name	Nazwa domyślnej bazy, jeśli nie jest ona nigdzie podana (brak w parametrach stron HTML).	3
DebugLevel	Zwykle powinien wynosić 0. Im większa jest wartość tego parametru, tym bardziej szczegółowy będzie poziom monitoringu systemu. Zaleca się ustawianie tego parametru powyżej 0 tylko w czasie testów. Wartość 20000 oznacza bardzo szczegółowy monitoring systemu.	4
DebugLogFile	Podawana jest nazwa katalogu (bezwzględna) oraz nazwa zbioru, gdzie zapisywane będą informacje nt. funkcjonowania systemu (tylko jeśli DebugLevel >0).	4
DefaultFormat	Jest to domyślny format wyświetlania danych, w przypadku, gdy żaden format nie zostanie wyspecyfikowany (w historii wyszukiwań lub gdy dane połączenie nie określa wartości format_name).	1
EntryDefinitionFile	Dłyczy zbioru interfejsu wprowadzania danych (entry.def).	1
FieldAndOr	Dla FieldAndOr równego 1 wyświetlone zostaną przyciski typu radio (pozwalające na łączenie zawartości pól operatorami Boolowskimi: AND, OR, NOT); w przeciwnym przypadku nie zostaną wyświetlone przyciski a pola będą łączone domyślnie operatorami AND.	6
Format	Dla wartości 1, ekran wyszukiwania będzie wyświetlał pole typu combo z możliwością wyboru formatu wyświetlania danych. W przeciwnym przypadku (=0) będzie wykorzystany format domyślny (pierwszy na liście lub wskazany przez parametr DefaultFormat).	6
HighlightSearchTerms	Ustawienie równe 1 spowoduje, że terminy wyszukiawcze będą zaznaczone w wynikach wyszukiawczych (kolorem czerwonym).	4
HtpFilePath	Jest to ścieżka dostępu do katalogu, w którym przechowywane są wersje językowe zbiorów HTTP. Zwykle zbiory są przechowywane w podkatalogu aplikacji. Pamiętajmy, że katalog http zawiera różne wersje językowe szablonów HTML, służących do budowania formularzy html danego interfejsu. Tak więc jeśli został określony parametr lang=eng, to system będzie poszukiwał zbiorów, które mieszczą się w kataloguhttp/eng/.	1

IncludeFilePath	Podobnie jak poprzedni parameter, ten dotyczy lokalizacji zbiorów języka JavaScript (o rozszerzeniu js), zwykle wskazujehttp/eng/	1
IsisDataDir	Parametr podaje ścieżkę dostępu do zbiorów baz danych ISIS. Ścieżka ta jest wykorzystywana za każdym razem ilekroć system odwołuje się do ISISDLL. Jeśli jednak parametr nie jest użyty lub też jest nieokreślony, zbiór ISISDLL próbuje pobrać dane ze zbioru dbn.par.	1
IsisDocLocationOcc	Ustala maksymalną liczbę wystąpień wirtualnego pola V9981	6
IsisDocLocation	Parametr ten jest wykorzystywany do ustalenia katalogu roboczego w przypadku odwoływania się do tzw. wirtualnego pola V9981, aby wskazać na zbiór dokumentów. Może być również wykorzystywany w celu budowania połączeń do dokumentów dołączonych do rekordów ISIS. Pożyteczny dla aplikacji WWW-ISIS na CD-ROM w czasie procedury instalacyjnej (inicjacyjnej) automatycznie ustala położenie zbiorów.	6
IsisForceType	Jeśli parametr ten równa się 1, program (skrypt cgi) wprowadza do wszystkich zapytań wartość T=type.	1
DBNFile	Wskazuje plik opisujący pełne ścieżki do baz danych wykorzystywanych w procesie wyszukiwania i wprowadzania danych.	3
IsisMultiDbFile	Wskazuje plik opisujący pełne ścieżki do baz danych w przypadku wyszukiwania jednoczesnego w wielu bazach danych (parametr ISISMULTIDB jest ustawiony na 1). W przypadku, gdy stosowane jest wyszukiwanie jednoczesne w kilku bazach danych parametr DBNFile jest ignorowany.	3
IsisMultiDbUser	Gdy parametr ten przyjmie wartość 1 umożliwia ustalenie przez użytkownika, które bazy będą przeszukiwane równocześnie. Gdy parametr ten przyjmie wartość 0 użytkownik nie może ustalić, które bazy będą przez niego przeszukiwane automatycznie, zamiast tego przeszukiwanie jest wykonywane na zbiorze baz zdefiniowanym przez administratora w zbiorze IsisMultiDbFile.	3
IsisMultiDb	Gdy parametr ten przyjmie wartość 1 umożliwia jednoczesne wyszukiwanie w wielu bazach danych.	3
IsisTempFileLife	Określa minimalny czas przechowywania tymczasowych zbiorów dla zapytań (w sekundach). Np. dla IsisTempFileLife równego 100, zapytania, które będą trwały dłużej niż 100 sekund w praktyce nie będą obsługiwane. W ustaleniu tej wartości należy kierować się możliwościami serwera WWW, a przede wszystkim natężeniem ruchu w sieci. Domyślna wartość jest ustalona na 600 (10 minut).	1
IsisWorkDir	Katalog, w którym ISISDLL tworzy tymczasowe zbiory. Pliki te są usuwane po (udanym) wykonaniu zlecenia.	1
LatinConvOn	Jeśli baza danych jest zapisana w standardzie DOS a chcemy aby znaki diakrytyczne były widoczne w naszej przeglądarce, parametr ten powinien wynosić 1. Jeśli zaś nasza baza została „przetłumaczona” zgodnie ze standardem kodowania znaków ISO 8859-1 lub ISO 8859-2, to winniśmy zmienić ustawienie na 0.	4
LatinConvTerms	Parametr ten informuje program cgi, czy nasze zapytanie winno być tłumaczone na ISO8859-x (1-tak; 0-nie). Pamiętajmy, aby parametr ustalić zgodnie z poprzednim LatinConvOn oraz konwersją ISISUC. Jeśli ISISUC usuwa znaki diakrytyczne, to nie jest wymagana konwersja.	4
ListCheckBox	Dla ListCheckBox= 1 wyświetlony będzie słownik wraz z przyciskami typu „zaznacz” (check). W przypadku ListCheckBox równego 0 na ekranie wyświetlone będzie pole typu combo z 20 terminami.	6
PagingMaxPages	Wskazuje jak wiele odwołań będzie wyświetlonych na stronie. Np. dla PagingMaxPages= 10 pokazywanych będzie połączenie od 1 do 10 lub od 11 do 20, 21 do 30 itd.	3
PagingMaxRows	Wskazuje, ile rekordów będzie wyświetlonych na stronie	3

QueryMaxRows	Parametr ten definiuje jak wiele rekordów będzie posortowanych i wyświetlonych jako rezultat wyszukiwania. Wielkość ta zależy w dużej mierze od możliwości obliczeniowych komputera, na którym zainstalowany jest serwer. Rozsądną wartością jest ustalanie tego parametru między 100 a 500 (pamiętajmy, że jest mało prawdopodobne, aby użytkownik chciał przeglądać więcej niż 500 rekordów, z drugiej strony parametr zabezpiecza przed pytaniami, które mają na celu ściągnięcie całej bazy danych)	4
RestrictedTermMinLength	Parametr ten definiuje minimalną długość maskowanego terminu wyszukiwawczego w znakach, konieczną do wprowadzenia w polach wyszukiwawczych (określonych parametrem <code>Restricted_term</code> w zbiorze <code>database.def</code>); zwykle jest to wartość między 2 a 3 znaki. Ma to na celu uniknięcia pytań typu <code>a\$</code>	3
SearchDefinitionFile	Jest to katalog, w którym przechowywany jest zbiór <code>database.def</code>	1
SearchPaging	Gdy parametr ten wynosi 1, rezultaty wyszukiwań będą podzielone na strony, zaś rozmiar strony jest zdefiniowany przez <code>PagingMaxRows</code> .	3
Sort	Ustawienie tej wartości równe 1 spowoduje, że wyniki będą mogły być posortowane zgodnie z kryteriami wybranymi z pola typu <code>combo</code> . W przeciwnym przypadku (<code>=0</code>) nie pojawi się możliwość sortowania wyników.	6
TermSeparator	Znaki oddzielające terminy wewnątrz pola wprowadzania danych, znaki te będą zinterpretowane jako operatory <code>AND</code> lub <code>OR</code> w zależności od ustawień dotyczących operatorów <code>AND/OR</code> . Standardowe ustawienia to <code>+</code> oraz <code>*</code> . UWAGA: Nie wykorzystuj pustego znaku jako separatora.	6



DODATKOWE UWAGI OGÓLNE:

- W początkowej fazie testowania systemu zaleca się ograniczyć zmiany pliku konfiguracyjnego wyłącznie do zmian parametrów obligatoryjnych, których ważność została określona jako 1, 3, a następnie 2.
- Powtórnie sprawdzić parametr (obligatoryjny), w szczególności sprawdzić ścieżki dostępu. Ich wadliwe ustawienie może spowodować problemy w pracy systemu.
- Jeśli nie jesteśmy pewni znaczenia jakiegoś parametru, to nie zaleca się jego zmiany. Dokładne informacje na temat parametrów znajdzie czytelnik w dokumentacji systemu.



USTAWIENIA `HttpFilePath`

Aby upewnić się, że korzystamy z właściwych szablonów HTTP i tym samym dokonać właściwych zmian, należy sprawdzić parametr odpowiedzialny za ustawienia języka czyli zmiennej `lang`.



Przykład 1

Jeśli parametr `httpFilePath` wynosi:

`c:\XITAMI\WEB\demo\http`

zaś parametr odpowiedzialny za wybór języka wynosi `eng`, to system będzie próbował znaleźć następujący katalog:

`c:\XITAMI\WEB\demo\http\eng`

PARAMETR IsisForceType



Parametr ten powinien być zastosowany, jeśli w naszej bazie danych przechowujemy więcej niż jeden typ rekordu, dzięki któremu możemy dokonać rozróżnienia w tworzonych zapytaniach. W takim przypadku niezbędne jest też odpowiednie przygotowanie zbioru FST, tak aby prefiks „T=” mógł utworzyć tzw. punkty dostępu przez wskazanie na typy rekordu występujące w bazie.

PARAMETR IsisDataDir



W systemie ISIS zwykle wskazujemy poszczególne pliki bazy danych za pomocą pliku dbn.par. W chwili obecnej ISISDLL odczytuje poprawnie dbn.par tylko w przypadku zbiorów głównego i odwróconego. Dlatego nie zaleca się przekierowywania ISISDLL za pomocą dbn.par. Zarówno pliki bazy danych, jak i formaty odwołujące do innych baz danych powinny być przechowywane w katalogu wskazanym przez IsisDataDir (o ile nie ma wskazania ścieżek do plików baz danych za pomocą parametru DBNFile).

PARAMETR IsisWorkDir



W pewnych przypadkach system operacyjny ignoruje ten parametr i zapisuje tymczasowe zbiory do katalogu temp (co może zależeć od ustawień zmiennych środowiskowych systemu operacyjnego). Dlatego zaleca się sprawdzenie, czy na dysku twardym pozostają zbędne pliki tymczasowe z rozszerzeniami: *.tph oraz *.tpq. Jeśli tak, to należy ustalić nazwę katalogu dla parametru IsisWorkDir równą lokalizacji tych plików, tak aby system był w stanie zlikwidować je po upływie maksymalnego czasu przechowywania plików tymczasowych (wskazanych przez IsisTempFileLife).

Aneks B

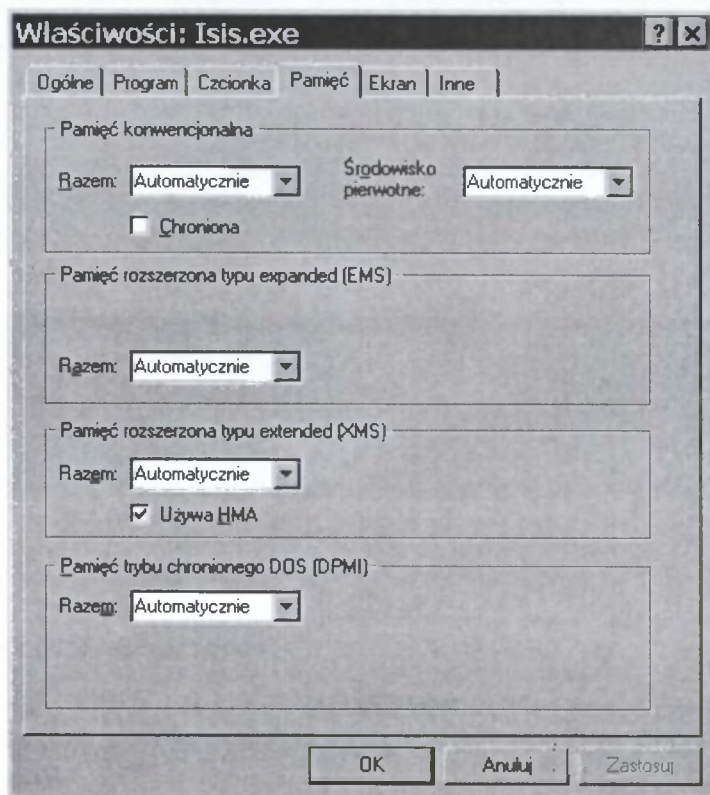
Elementy systemu CDS/ISIS wyciąg z dokumentacji

Poniższy aneks został umieszczony głównie ze względu na wygodę czytelnika – w celu zwiększenia kompletności omawianego materiału i stanowi w dużym stopniu wyciąg z dokumentacji systemu. Omówione są pokrótce najistotniejsze aspekty systemu ISIS niezbędne dla zrozumienia funkcjonowania pakietu WWW-ISIS, jego wdrożenia i eksploatacji. W aneksie zostały przedstawione podstawowe funkcje języka wyszukiwawczego CDS/ISIS oraz języka formatowania. Omówione zostało w zarysie zagadnienie tworzenia baz danych w CDS/ISIS, techniki tworzenia zbiorów odwróconych. Pragniemy jeszcze raz podkreślić, że aneks ten nie stanowi próby wyczerpującego omówienia systemu CDS/ISIS. W przypadku jakichkolwiek wątpliwości zalecane jest zapoznanie się z pełną dokumentacją systemu CDS/ISIS oraz ze stosownymi publikacjami nt. CDS/ISIS.



UWAGA:

Zaleca się wykorzystanie wersji ISIS dla systemu operacyjnego DOS i uruchamianie go jako aplikacji DOS. We właściwościach programu należy upewnić się, czy są poprawne ustawienia pamięci systemu – dla wszystkich opcji zaleca się wybór automatyczny (rys. 77).



Rys. 77. Ustawienia pamięci dla programu ISIS (wersja dla DOS)

Dostępność oprogramowania i pomocy technicznej

Od dłuższego czasu UNESCO dystrybuuje kopie CDS/ISIS przez ponad 138 akredytowane ośrodki w różnych krajach na całym świecie. Warunki korzystania z oprogramowania CDS/ISIS są określone w umowie podpisanej między UNESCO (lub oficjalnym dystrybutorem), a instytucją otrzymującą oprogramowanie.

ISIS nie jest produktem komercyjnym, zatem użytkownicy nie mogą liczyć na typowe dla produktów komercyjnych wsparcie techniczne. Korzystanie z systemu polega na podejściu „zrób to sam” i zmusza użytkownika do poszukiwania we własnym zakresie rozwiązań. Istnieją jednak różne sposoby uzyskania wsparcia technicznego:

- W wielu krajach istnieją specjalne grupy użytkowników, dzięki którym można uzyskać pomoc.
- Wiele grup użytkowników organizuje kursy na temat korzystania z ISIS na różnych poziomach zaawansowania.
- w Internecie istnieje szereg grup dyskusyjnych (np. <http://www.agralin.nl/isis/isislist.html>).
- Istnieje wiele witryn internetowych, na których można znaleźć szereg użytecznych informacji związanych z bazami ISIS, jak: przykładowe aplikacje, bazy oraz narzędzia do rozwiązywania szeregu problemów informatycznych (np. <ftp.unesco.org/pub/ftp.bireme.br>, <http://www.agralin.nl/isis> itd.).

Struktura baz danych i architektura systemu

Bazy CDS/ISIS

Podstawowe elementy danych przechowywane są w polach. Każde pole posiada niepowtarzalny znacznik numeryczny. Znacznik ten pełni rolę zbliżoną do nazwy pola. Znaczniki pól przyporządkowywane są poszczególnym polom w trakcie procesu definiowania struktury bazy. Na przykład w trakcie definiowania struktury bazy bibliograficznej możemy przyporządkować polu „autor” znacznik 10, zaś polu „tytuł” znacznik 20 itd.

Jednakże może się okazać, że pole będzie zawierało więcej niż jeden element danych. W takich przypadkach pole może być podzielone na dwa lub więcej podpól. Podpola są definiowane dzięki tzw. ogranicznikom podpól. Ograniczniki te są dwuznakowymi kodami: pierwszy znak to „^”, zaś drugi to dowolny znak alfanumeryczny (np. a, b, c, itd.), zatem znacznik podpola może przyjąć postać „^a”. Pole może być opcjonalne (czyli nie musi występować w każdym rekordzie) lub powtarzalne (dany rekord może zawierać więcej niż jedno wystąpienie, np. pole „deskryptory” zwykle zawiera więcej niż jedno słowo kluczowe opisujące zawartość dokumentu).

CDS/ISIS został zaprojektowany tak, aby przetwarzać dane w polach o różnej długości. Znaczy to, że system będzie przechowywał tylko tyle znaków, ile rzeczywiście zostało wprowadzonych do danych rekordów¹ (np. 20, 30, 50... itd.). Cecha ta pozwala na optymalne wykorzystanie powierzchni dysku i zapewnia daleko idącą elastyczność w ustalaniu maksymalnej długości pola.

Zbiór konfiguracyjny syspar.par

W celu umożliwienia korzystania z systemu konieczne jest poprawne skonfigurowanie zbioru konfiguracyjnego, w którym określone jest umiejscowienie plików systemowych oraz plików użytkownika. Przykładowy zbiór parametrów zbioru syspar.par wygląda następująco:

¹ W tablicy Definicji Pól (FDT) definiuje się co prawda maksymalną długość pola, ale ma to tylko znaczenie dokumentacyjne.

1=PROG\
2=MENU\
3=MSG\
4=WORK\
5=Auth\

System przewiduje ponad 10 parametrów, jednak do poprawnego funkcjonowania pakietu CDS/ISIS wystarcza właściwe skonfigurowanie 5. Poniżej zostaną omówione w skrócie niektóre z nich:

Parametr **0** – wskazanie innego zbioru parametrów, który nadpisuje parametry z Syspar.par²

Parametr **1** – ścieżka dostępu do plików programów

Parametr **2** – ścieżka dostępu do plików menu

Parametr **3** – ścieżka dostępu do plików komunikatów

Parametr **4** – ścieżka dostępu do plików roboczych

Parametr **5** – ścieżka dostępu do plików baz danych

Parametr **6** – nazwa domniemanej bazy danych

Parametr **7** – język domniemany

Parametr **8** – ogranicznik pól powtarzalnych

Funkcje systemu

Generalnie system CDS/ISIS daje możliwość:

- Definiowania baz zawierających elementy danych.
- Wprowadzania nowych rekordów do bazy.
- Modyfikacji, korekty lub usuwania istniejących rekordów.
- Automatycznego tworzenia i utrzymywanie zbiorów pozwalających na szybki dostęp do każdej bazy.
- Przeszukiwania bazy przy wykorzystaniu zaawansowanego języka wyszukiwawczego.
- Wyświetlania rekordów.
- Sortowania rekordów w żądanym porządku.
- Wydruk katalogów i/lub indeksów.
- Rozwój zaawansowanych aplikacji przy wykorzystaniu zintegrowanych narzędzi systemu (jak np. języka PASCAL).

Wszystkie wymienione środki programistyczne dostępne są dzięki 8 podstawowym usługom zgrupowanym w dwóch kategoriach: Usługi użytkownika – działających na istniejących bazach; Usługi systemowe – zaprojektowane dla administratora bazy w celu definiowania nowych baz i wykonywania różnorodnych zadań systemowych.

Usługi systemowe dotyczą następujących funkcji (podsystemów):

- **ISISENT** – wprowadzanie danych i edycja rekordów;
- **ISISRET** – wyszukiwanie danych;
- **ISISPRT** – wydruki;
- **ISISINV** – obsługa zbioru odwróconego.

Z kolei usługi systemowe dotyczą następujących funkcji

- **ISISDEF** – definiowanie struktury nowej bazy i modyfikacja istniejących;
- **ISISUTL** – funkcje narzędziowe;
- **ISISXCH** – funkcje archiwizacji i wymiany danych;
- **ISIPAS** – zaawansowane narzędzia programistyczne;

Struktura bazy

Mimo że baza CDS/ISIS przedstawia się użytkownikowi jako pojedynczy zbiór informacji, w zasadzie składa się ona z szeregu logicznie połączonych i fizycznie odrębnych zbiorów

² Technika ta w wersji DOS-owej zastępuje sprawdzanie hasła.

rów. Zarządzanie tymi zbiorami jest domeną samego systemu i użytkownik nie musi znać ich szczegółów, aby korzystać z bazy. Jednakże znajomość pewnych elementów dotyczących połączonych zbiorów może pozwolić lepiej zrozumieć funkcjonowanie systemu. Proponujemy przesłanie tablicy definicji pól bazy AGRIS (patrz Aneks F).

Definiowanie struktury bazy

Zanim rozpoczniemy przetwarzanie danych niezbędnym jest zdefiniowanie w CDS/ISIS struktury rekordu i jego zawartości, która jest określona w zbiorach definiujących bazę. Odpowiedzialna za to jest usługa ISISDEF, która pozwala na tworzenie i modyfikacje tych zbiorów.

Zbiory definiujące bazę	
Tablica Definicji Pól (Field Definition Table - FDT)	Definiuje pola, które będą występowały w bazie
Arkusze wprowadzania danych (Data entry worksheet(s))	Układ formularza (arkusza) do wprowadzania/uaktualniania danych (rekordów)
Formaty wyświetlania danych (Display format(s))	Wyrażenia definiujące sposób w jaki dane będą wyświetlane, drukowane lub pobierane
Tablica(e) Wyboru Pól (Field Select Table(s) (FST))	Definiuje zbiór przeszukiwanych pól w bazie, ponadto FST definiuje warunki sortowania i formatowania w bazie.

Choć można zdefiniować strukturę bazy korzystając z odpowiednich podsystemów systemu CDS/ISIS, równie dobrze można zapisać całą informację na temat bazy za pomocą edytora tekstu w zbiorze, którego fragment przedstawia tabela 20.

Tabela 20

Fragment tablicy definicji pól w CDS/ISIS.

Tytuł pola (pojawia się w formularzach CDS/ISIS)	Ograniczniki podpól (jeśli występują)	Znaczniki pola (maks)	Długość pola	0 -alfanumeryczne; 1- alfabetyczne; 2- numeryczne; 3-wg wzorca 1 -powtarzalne; 0 -niepowtarzalne	
***	***	***	***	*	*
second. subj. cat.		9	7	0	1
type of document		10	1	1	0
bibliogr. levels		11	3	1	0
literary indicator		12	10	1	1
***	***	***	*	*	*
Language	^a^b^c	500	100	0	1
literary indicator -Extension		13	1	1	0
corporate name(s)		110	400	0	0
degree		111	80	0	0
***	***	***	***	*	*

Zbiór główny

Zbiór główny (plik o rozszerzeniu mst) zawiera wszystkie rekordy danej bazy. Każdy rekord jest identyfikowalny przez niepowtarzalny numer automatycznie przyporządkowany przez system CDS/ISIS w procesie tworzenia rekordu, nazywany Master File Number (MFN).

Aby umożliwić szybki dostęp do dowolnego rekordu, system CDS/ISIS tworzy specjalny zbiór (o rozszerzeniu xrf) – będący skorowidzem zbioru głównego. Relacja między zbiorem głównym a skorowidzem została przedstawiona w rozdziale poświęconym systemowi ISIS (rozdz. 2).

Tworzenie, utrzymywanie oraz modyfikacja zbioru głównego są wykonywane przez funkcje wprowadzania danych – podsystem ISISENT.

Zbiór odwrócony

Rekordy mogą być wyszukiwane bezpośrednio przez podanie ich numeru MFN. Aby umożliwić dodatkowe punkty dostępu do rekordów, system CDS/ISIS tworzy tzw. zbiór odwrócony, który de facto jest indeksem całej zawartości bazy i umożliwia szybki dostęp do rekordów przez np. podanie autora, tytułu itd.

Zbiór odwrócony zawiera wszystkie terminy, które mogą być wykorzystane jako punkty dostępu podczas procesu wyszukiwania, zaś dla każdego z nich zostanie również podana liczba odwołań do rekordów, w których występują te terminy. Zbiór wszystkich punktów dostępu dla danej bazy nazywany jest zbiorem odwróconym, słownikiem lub indeksem. Dlatego też zbiór odwrócony można również postrzegać jako indeks zawartości zbioru głównego.

Na przykład cztery rekordy ze zbioru odwróconego (z numerem 18, 204, 766 oraz 1039) zawierają słowa KSZTAŁCENIE MŁODZIEŻY. Logiczna struktura wycinka stosownego zbioru odwróconego byłaby następująca: KSZTAŁCENIE MŁODZIEŻY 18 204 766 1039.

Jak widać KSZTAŁCENIE MŁODZIEŻY jest punktem dostępu (lub terminem słownika) z odwołaniami do rekordów ze zbioru głównego. Ponieważ każdy termin ma przeważnie inną liczbę odwołań do rekordów ze zbioru głównego, logiczne rekordy w zbiorze odwróconym są zmiennej długości. Podobnie, jak w przypadku zbioru głównego zbiór odwrócony składa się fizycznie z szeregu powiązanych ze sobą zbiorów.

Za pomocą tablicy wyboru pól (Field Select Table – FST) określa się technikę selekcji punktów dostępu z pól rekordów zbioru głównego, które mają znaleźć się w zbiorze odwróconym. Inne systemy również wykorzystują technikę zbiorów odwróconych, lecz tworzony jest osobny zbiór odwrócony (indeks) dla każdego (przeszukiwanego) pola (atrybutu). CDS/ISIS wykorzystuje pojedynczy (logicznie) zbiór dla danej bazy. Można założyć, że ze względu na swoją funkcjonalność zbiór odwrócony jest ekwiwalentem wielu zbiorów odwróconych (indeksów) w innych systemach. Obecna wersja CDS/ISIS skraca wszystkie terminy w zbiorze odwróconym do 30 znaków.

W momencie określania Tablicy Selekcji Pól służącej do budowy zbioru odwróconego konieczna jest znajomość języka formatowania, dzięki któremu dokonujemy wyboru pola lub jego fragmentu do tworzonego indeksu. Jednak pozostaje jeszcze tzw. technika selekcji. Dzięki technice selekcji określone zostaną reguły tworzenia punktów dostępu z tekstu pobranego z rekordu zgodnie z formatem. Istnieje 9 technik selekcji:

- Technika selekcji **0** – jeżeli sformatowany tekst składa się z kilku wierszy, z każdego wiersza budowany jest punkt dostępu; technikę tę stosuje się zwykle do wyboru do indeksu pełnych pól i podpól;
- Technika selekcji **1** – z każdego wiersza wybierane są podpola jako punkty dostępu; technika ta tworzy indeks z całych podpól;
- Technika selekcji **2** – punkt dostępu budowany jest z każdego terminu lub frazy ujętej w nawiasy ostre ('<.....>');

Technika selekcji 3 – punkt dostępu budowany jest z każdego terminu lub frazy ujętej w ukośniki ('/...../');

Technika selekcji 4 – punkt dostępu budowany jest z każdego słowa pobranego z tekstu przy użyciu formatu.

Wraz z wersją systemu 3.07 zostały wprowadzone 4 dodatkowe techniki selekcji (5, 6, 7, 8) pozwalające na podanie prefiksu dla każdego wygenerowanego punktu dostępu. Odpowiadają one odpowiednio technikom: 1, 2, 3 i 4.

Prefiks jest określony w formacie ekstrakcji danych jako literal bezwarunkowy według podanego schematu:

`'dp...pd', [format]`

d – ustalony ogranicznik pola (nie występujący w prefiksie)

p – bieżący prefiks

Rozważmy specyfikację z techniką selekcji poniżej:

`1 8 '/T=/',v24`

Oznacza ona, że każde słowo wyselekcjonowane z pola 24 i poprzedzone prefiksem T= będzie punktem dostępu w zbiorze odwróconym.

Zbiory ANY

Dodatkowym i uzupełniającym w stosunku do zbioru odwróconego, choć nie zawsze występującym typem zbioru, jest zbiór ANY. Jest on używany w celu połączenia razem powiązanych ze sobą znaczeniowo terminów. Termin należący do grupy ANY jest zbiorową nazwą przyporządkowaną grupie terminów wyszukiwawczych. Kiedy używany jest termin ANY, wywoływana jest cała lista terminów, zaś terminy te są automatycznie łączone operatorem sumy logicznej (Boolowskie OR).

Tworzenie i dobór terminów ANY zależy w dużej mierze od częstotliwości ich występowania w zapytaniach wyszukiwawczych. Najlepszymi kandydatami są nazwy geograficzne. Jeśli na przykład stworzymy termin ANY dla wyrażenia Ameryka Łacińska (Meksyk, Peru, Ekwador), to w przyszłości będziemy mogli używać tej zbiorowej nazwy w procesie wyszukiwania, zamiast wpisywać za każdym razem wszystkie nazwy krajów Ameryki Łacińskiej.

Relacje między zbiorami

Logiczne relacje między podstawowymi zbiorami w bazie CDS/ISIS są lepiej zrozumiałe, gdy dokładniej przyjrzymy się procesowi wyszukiwawczemu. Wyszukiwanie danych w bazie jest wykonywane przez odwołanie się do zbioru odwróconego i wyszukanie tam terminów użytych w zapytaniu, w konsekwencji odnajdywane są numery rekordów (MFN), w których te terminy występują. Wyszukane w zbiorze odwróconym listy rekordów są przetwarzane przez program w zależności od zastosowanych w zapytaniu operatorów wyszukiwawczych. Pod koniec całego procesu utworzona jest ostateczna lista rekordów, które spełniają zadane w zapytaniu kryteria wyszukiwawcze, po czym rekordy z listy są wyświetlane zgodnie z ustalonym porządkiem.

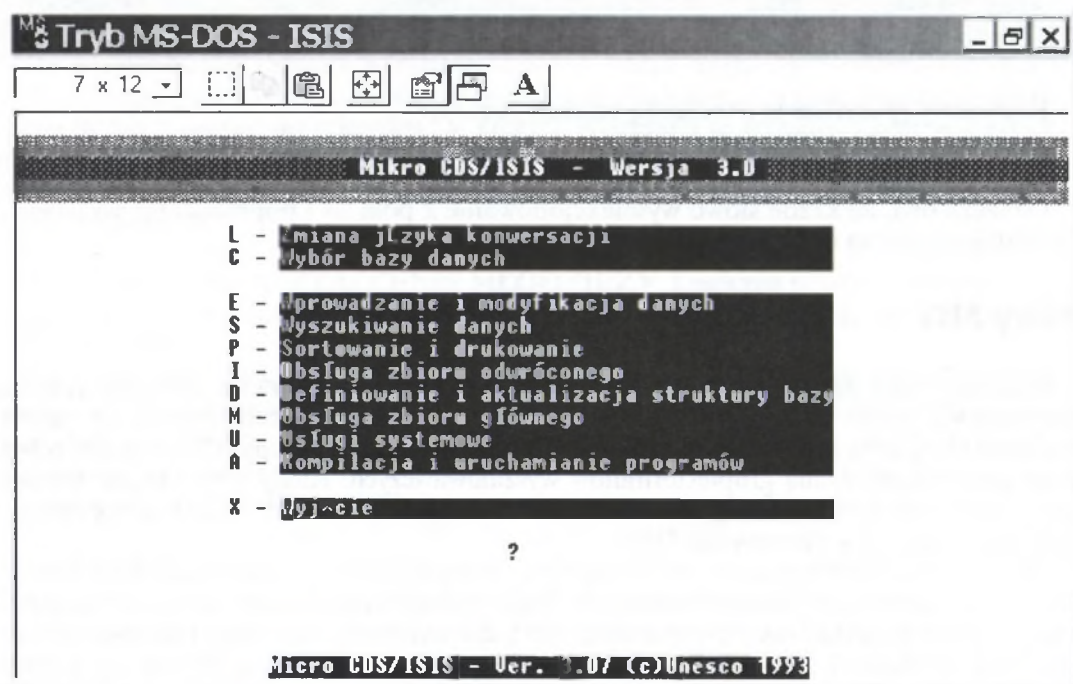
Użytkownik może zachować utworzoną listę wyników lub wydrukować jej zawartość dzięki podsystemowi ISISPRT.

Architektura systemu

Funkcje systemu CDS/ISIS dostępne są z połączonych ze sobą systemów menu. Jednakże, aby w pełni efektywnie zarządzać bazami, należy zapoznać się również z licznymi dodatkowymi technikami, specyficznymi dla CDS/ISIS, takimi jak język wyszukiwawczy czy język formatowania.

Systemy Menu

Większość koniecznych funkcji systemu dostępna jest z pozycji hierarchicznie zorganizowanego systemu menu. Menu jest listą dostępnych funkcji systemu. Użytkownik wprowadza pojedynczy znak z klawiatury (dla danej wersji językowej) i wywołuje daną usługę np. dla wersji angielskiej systemu ISIS wybranie „E” z menu głównego spowoduje wywołanie funkcji podsystemu związanych z obsługą wprowadzania i modyfikacji danych (menu ISISENT).

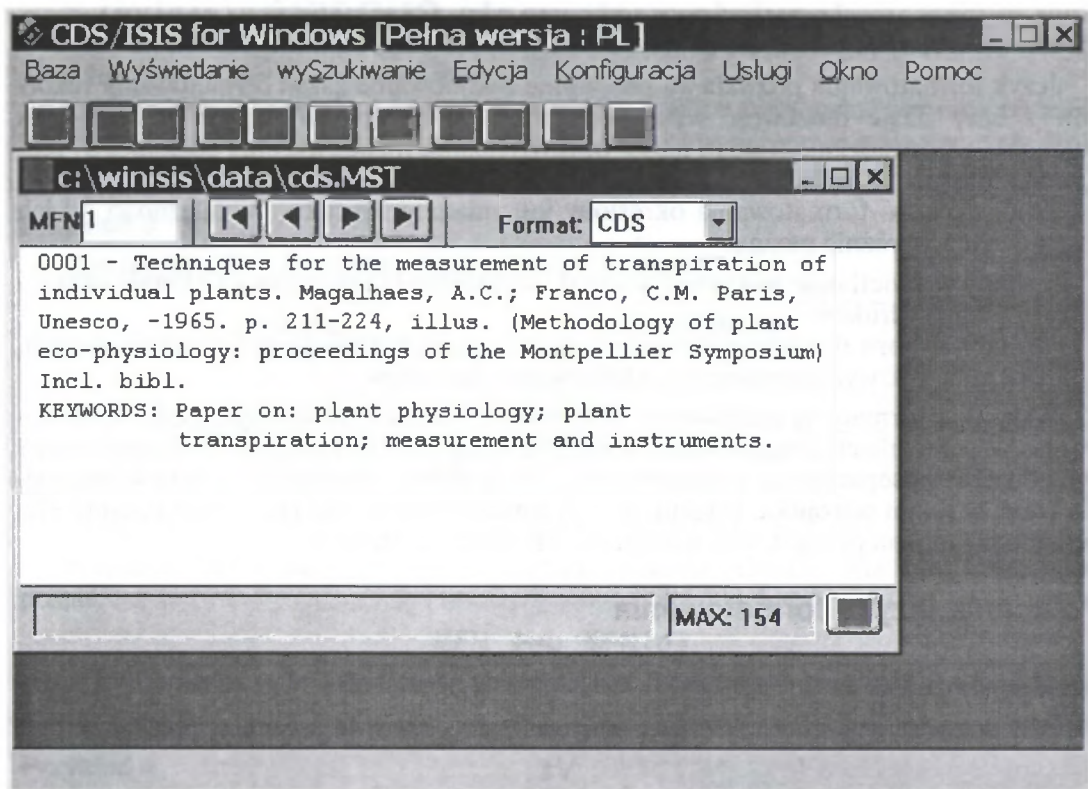


Rys. 78. Przykładowe menu CDS/ISIS dla DOS

Po wybraniu danej opcji z systemu menu, użytkownik proszony jest o wprowadzenie dodatkowych informacji, np. nazwy bazy danych. System wyświetla żądanie na dole ekranu, po wprowadzeniu informacji należy przycisnąć klawisz Enter.

Po otwarciu bazy każdy kolejny podsystem wyświetla status nowo otwartej bazy w dwóch ostatnich wierszach na dole ekranu:

- Nazwę bazy.
- Nazwę bieżącego arkusza wprowadzania danych.
- Nazwę bieżącego arkusza wyświetlania danych.
- Ostatni numer MFN (ostatni numer rekordu).



Rys. 79. Przykładowe menu CDS/ISIS dla Windows

Wielojęzyczność systemu

System CDS/ISIS został tak zaimplementowany, aby można było łatwo stworzyć dowolną wersję językową i przełączać się podczas pracy z jednej wersji językowej na inną. W tym sensie system CDS/ISIS jest wielojęzyczny. Na przykład, dla DOS-owej wersji angielskiej zmiana języka komunikacji z systemem może być przeprowadzona przez wybór opcji L – Change dialogue language w głównym menu.

Wielojęzyczność systemu jest ograniczona jedynie możliwościami reprezentacji znaków alfabetu konkretnego systemu operacyjnego.

Arkusze wprowadzania danych

Arkusze ten jest wyświetlany, gdy użytkownik ma zamiar wprowadzić dane. Składa się on z odpowiedniego zbioru pól, z których każde posiada nazwę oraz puste miejsca do wprowadzania danych. W pewnych przypadkach pole może już zawierać dane jako wartości domyślne.

Jeden arkusz może zawierać więcej niż jedną, lecz nie więcej niż 19 stron. Istnieją dwa typy arkuszy:

- Arkusze systemowe: używane przez CDS/ISIS zamiast tzw. zachęty systemowej.
- Arkusze wprowadzania danych: używane przez CDS/ISIS w celu tworzenia i modyfikacji rekordu. Arkusze tego typu mogą być definiowane przy wykorzystaniu wbudowanego edytora arkuszy. Dla jednej bazy można zdefiniować więcej niż jeden arkusz wprowadzania danych.

Język formatowania CDS/ISIS

Język formatowania pozwala na precyzyjne definiowanie zasad formatowania rekordów z bazy. Daje możliwość wyboru określonych elementów danych w dowolnym porządku, wprowadzanie stałego tekstu oraz określania rozmieszczenia poziomego i pionowego tekstu.

Zbiór poleceń formatowania określony jest mianem formatu. W zależności od ich zastosowania wyróżnić można:

- Formaty wyświetlania: wykorzystywane do wyświetlania rekordów na ekranie oraz do tworzenia wydruków.
- Formaty wyboru danych: wykorzystywane w tablicach FST do definiowania danych, które mają być wykorzystane do indeksowania rekordów.

Wszystkie formaty są definiowane przy wykorzystaniu tego samego języka formatowania. W poleceniach formatowania rozróżnialne są małe i duże litery. Polecenia winny być od siebie odseparowane przecinkiem (",") i są wykonywane przez system sekwencyjnie (czyli w takim porządku, w jakim zostały umieszczone w zbiorze formatującym). Dla jednej bazy można przygotować nieograniczoną liczbę formatów.

Polecenia języka formatowania

Wybór pola

Aby dokonać wyboru konkretnych wartości z pola rekordu, używamy polecenia:

Vt

gdzie V jest mnemotechnicznym oznaczeniem samej funkcji, zaś t jest znacznikiem danego pola zgodnie z tablicą definicji pól (FDT).

Np.:

Znacznik pola: 1200

Zawartość pola: Wstęp do Projektowania Systemów Informacyjnych

Format: v1200

Wynik działania formatu: Wstęp do Projektowania Systemów Informacyjnych

Wybór podpola (lub jego fragmentu)

Aby dokonać selekcji wartości z podpola rekordu, używamy polecenia:

Vt^i

gdzie V jest mnemotechnicznym oznaczeniem samej funkcji, t jest znacznikiem danego pola, zaś i jest znacznikiem podpola.

Np.:

Znacznik pola: 1100

Ograniczniki podpól: ^a i ^b

Przykład 2

Zawartość pola: ^aWarszawa^b1998

Format: v1100^a

Wynik działania formatu: Warszawa

Wybór fragmentu pola lub podpola

W pewnych przypadkach, zwłaszcza kiedy pole lub podpole ma taki sam układ w całej bazie danych (np. data), może zaistnieć potrzeba wyodrębnienia jego części. Służą temu polecenia przesunięcia, w następującej postaci:

*** przesunięcie.długość**

*** (gwiazdka)** przesunięcie – wskazuje pozycję pierwszego znaku, który ma zostać wybrany z pola lub podpola (pozycje znaków liczone są od zera, gdy parametr ten jest pominięty wybierane są wszystkie znaki);

• (kropka) długość – wskazuje ilość znaków, które mają zostać wybrane od znaku wskazanego przez przesunięcie, gdy parametr ten jest pominięty zostają wyprowadzone wszystkie znaki od pozycji wskazanej przez przesunięcie.

Przykład 3

Zawartość pola o wyróżniku (znaczniku) 1: 98-maj-07

Format	Tekst wynikowy
V1*3.3	maj
V1.2	98
V1*7, V1*2.4,	07-maj

Polecenia wyprowadzenia numeru rekordu (MFN)

Polecenie MFN służy do wyprowadzenia numeru rekordu. Ma ono następującą postać:

MFN lub MFN(*d*)

gdzie *d* jest liczbą cyfr, które będą wyświetlone. Jeśli *d* jest pominięte, wyświetlonych zostanie 6 cyfr.

Przykład 4

Np. dla rekordu o numerze 111 możemy mieć następujące przypadki:

Polecenie: MFN

Wynik: 000111

Polecenie: MFN(4)

Wynik: 0111

Polecenie: MFN(2)

Wynik: 11

Oczywiście w ostatnim przypadku wynik nie wskazuje poprawnie numeru rekordu.

Polecenia wyboru trybu

CDS/ISIS może wyświetlać dane w trzech różnych trybach:

- Tryb kontrolny: zawartość pól wyświetlana jest dokładnie w takiej kolejności i postaci, w jakiej zostały wprowadzone do bazy.
- Tryb nagłówkowy: wszystkie znaki kontrolne i sterujące umieszczone wewnątrz danych (np. <> itd) są ignorowane, zaś ograniczniki podpól są zamieniane na znaki interpunkcji.
 - ^ a zastępowany jest średnikiem -';
 - ^ b do ^ i zastępowane są przecinkiem - ','
 - pozostałe ograniczniki zastępowane są kropką - '.'
- Tryb danych – jest podobny do trybu nagłówkowego z tą różnicą, że każde pole jest kończone kropką, po której następują dwie spacje (jeśli pole kończy się kropką, dodawane są tylko dwie spacje).

Ponadto istnieje możliwość zamiany danych na duże litery.

Polecenia dotyczące trybów mają następującą postać:

Mmc

m określa tryb:

- P – tryb kontrolny
- H – tryb nagłówkowy
- D – tryb danych

c oznacza zamianę na wielkie/małe litery:

- U dane są zamieniane na duże litery
- L dane są pozostawiane bez zmian

Tak więc tryby wyświetlania danych mogą przyjmować następujące konfiguracje:

- MPL – tryb kontrolny
- MPU – tryb kontrolny z zamianą na wielkie litery
- MHL – tryb nagłówkowy
- MHU – tryb nagłówkowy z zamianą na wielkie litery
- MDL – tryb danych
- MDU – tryb danych z zamianą na wielkie litery

W jednym formacie można zastosować dowolną ilość poleceń. Domyślny jest tryb kontrolny bez zamiany małych litera na duże (czyli MPL).

Przykład 5

Zawartość pola: ^ aTalbott ^ bMary Ann

Tryb	Wynik
mpl	^ aTalbott^bMary Ann
mhu	TALBOTT, MARY ANN
mdl	Talbott, Mary Ann.



UWAGA:

Zmiana na duże litery odbywa się przez przekodowanie za pomocą tablicy ISISUC.

Polecenia rozmieszczenia pionowego i poziomego



UWAGA:

W przypadku WWW-ISIS nie zaleca się korzystania z poleceń rozmieszczenia pionowego i poziomego (w zastępstwie stosuje się odpowiednie znaczniki HTML jako literały).

Literały

Literał jest tekstem umieszczonym w odpowiednich ogranicznikach, który jako niezmienny jest wprowadzony do wyników danych formatu. **Istnieją trzy typy literałów:**

1. Literał **warunkowy**: zawarty w podwójnych znakach cudzysłowu " ", np. "Data wydania".
2. Literał **powtarzalny**: tekst zawarty między kreskami pionowymi | | np. |Autor:|.
3. Literał **bezwarunkowy**: tekst zawarty w pojedynczych znakach cudzysłowu ' ' np. 'Deskryptory:'

Literały warunkowe oraz powtarzalne są zawsze powiązane z polem lub podpolem przez swoje umiejscowienie w formacie: literały poprzedzające specyfikację pola (zwane preliterałami) są wyprowadzane przed zawartością związanego z nimi pola. Zaś literały

występujące po specyfikacji pola (zwane postliterałami) są wyprowadzane po zawartości związanego z nimi pola.

Jeśli bezpośrednio po preliterale warunkowym występuje znak + (np. |zzz|+), to literał będzie wyprowadzany przed wszystkimi wystąpieniami, z wyjątkiem pierwszego wystąpienia pola powtarzalnego.

Jeśli bezpośrednio przed postliterałem warunkowym występuje znak + (np. +|zzz|), to literał będzie wyprowadzany po wszystkich wystąpieniach, z wyjątkiem ostatniego wystąpienia pola powtarzalnego.

Literały bezwarunkowe są wyprowadzane zawsze niezależnie od występowania pola.

Pamiętajmy jednak, że jeżeli planujemy wykorzystanie WWW-ISIS, to w literałach można wykorzystać znaczniki HTML np.: ...""v10""... (czyli zawartość pola 10 będzie wyświetlona w przeglądarce pismem wytłuszczonym).

Przykład 6

Zawartość przykładowego rekordu:

Pole 10: ^aDonnelly^bJason

Pole 20: An Overview to Plant Physiology

Pole 80 w. 1: Plant Physiology

Pole 80 w. 2: Plant Sciences

Pole 80 w. 3: Biology

Format: mhl,v10^b" ",v10^a," - "v20

Wynik:

Jason Donnelly – An Overview to Plant Physiology

Format: mhl,"Author: "v10 / "Title: " v20/"#Descriptors:"v80+ |; |

Wynik:

Author: Donnelly, Jason Title: An Overview to Plant Physiology

Descriptors: Plant Physiology; Plant Sciences; Biology

Grupy powtarzalne

Grupy powtarzalne są zbiorem poleceń formatowania umieszczonym w nawiasach. Wszystkie polecenia umieszczone wewnątrz powtarzalnej grupy są wykonywane osobno dla każdego wystąpienia pola. Gdy nie jest używana grupa powtarzalna, wszystkie powtarzalne pola są traktowane jako pojedynczy ciąg znaków.

Grupy powtarzalne nie mogą być zagnieżdżane (tj. grupa powtarzalna nie może zawierać innej grupy powtarzalnej).

Przykład 7

Zawartość przykładowego rekordu:

Pole 10 w.1: ^aDonnelly^bJason

Pole 10 w.2: ^aTalbott^bMary Ann

Pole 20: An Overview to Plant Physiology

Pole 80: Plant Physiology%Plant Sciences%Biology

Format: v10^a| |,v10^b+|; |

Wynik: Donnelly Talbott Jason; Mary Ann

Format: (v10^a| |,v10^b+|; |)

Wynik: Donnelly Jason; Talbott Mary Ann

Format: v80/'****'

Wynik:

Plant PhysiologyPlant SciencesBiology

Format: (v80/)'****'

Wynik:

Plant Physiology

Plant Sciences

Biology

Funkcje numeryczne– funkcja val

Funkcja val zwraca wartość numeryczną swojego argumentu, który może zawierać dowolne poprawne wyrażenie języka formatowania.

Przykład 8

Zawartość pola o wyróżniku 10 wynosi 11.

Format val (v10) zwróci wartość "11".

Format val('111,100') zwróci wartość 111,100.

Funkcje wykonywane na łańcuchach znaków – funkcja REF

Funkcja REF pozwala na wybór danych z alternatywnego rekordu pliku głównego. Argument pierwszy jest wyrażeniem numerycznym i podaje numer rekordu (MFN), argument drugi określa format, który ma być zastosowany do danego rekordu. W przypadku kiedy wartość wyrażenia nie odpowiada żadnemu z numerów rekordu (MFN) istniejących w bazie, funkcja REF zwraca pusty łańcuch znaków.

Załóżmy, że mamy dwa rekordy:



UWAGA:

W WWW-ISIS polecenie REF nie działa poprawnie zamiast tego należy stosować ISISREF.

o numerze 1(MFN=1):

Pole 10: Abacki Jan

Pole 20: Zagadnienia sieci semantycznych

Pole 30: 110

o numerze 110(MFN=110):

Pole 10: Zagadnienia Informacji Naukowej

Pole 20: Instytut Informacji Naukowej

Pole 13: Warszawa 1990

Format: mpl v10/v20/ ref(val(v30),v11)

Tekst wynikowy:

Abacki Jan

Zagadnienia sieci semantycznych

Zagadnienia Informacji Naukowej



UWAGA:

Znak '/' jest poleceniem przejścia do następnego wiersza. W języku HTML przejście do nowej linii nie ma znaczenia w sposobie wyświetlania danych, dlatego w formatach dla WWW-ISIS, można go stosować jedynie dla poprawienia czytelności formy źródłowej HTML, nie zaś obrazu na ekranie przeglądarki.

Format: ref(val(v30),v13)

Tekst wynikowy:

Warszawa 1990

Instrukcja warunkowa

Instrukcje warunkowe typu „if ... then” pozwalają budować formaty uwzględniające kontekst (tj. takich formatów, których sposób formatowania zależy od zawartości formatowanego rekordu). Ogólna postać warunkowej instrukcji formatowania wygląda następująco:

if warunek then format_1 else format_2 fi

warunek jest wyrażeniem boolowskim;

format_1 jest formatem CDS/ISIS, który będzie wykonywany tylko wtedy, gdy wyrażenie Boolowskie przyjmie wartość prawdy;

format_2 jest formatem CDS/ISIS, który będzie wykonywany tylko wtedy, gdy wyrażenie Boolowskie przyjmie wartość fałszu.

Klauzula **else format_2** jest opcjonalna i może być pominięta. Słowa kluczowe **if**, **then** oraz **fi** są obowiązkowe.

Przykłady wyrażeń warunkowych mogą być następujące:

Vt : 'ciąg_znaków'

(pole o wyróżniku t zawierające wskazany ciąg znaków) – np. V80:'econom'

p(Vt)

(pole o wyróżniku t, które występuje w danym rekordzie) np. p(v10)

a(Vt)

(pole o wyróżniku t, które nie występuje w danym rekordzie) np. a(v20)

VAL(Vt) > wartość_numeryczna

(wartość z pola t jest większa niż wskazana wartość_numeryczna, można także używać innych operatorów matematycznych jak <, =, <>, <=, >= operator) – np. VAL(v50) > 1980

Możliwe jest budowanie kompleksowych wyrażeń Boolowskich przy wykorzystaniu operatorów: OR, AND, NOT oraz odpowiednio umieszczonych nawiasów np. (V80:'agricult' OR V80:'forest') AND P(V10)

Wyrażenia **if** można zagnieżdżać do dowolnego poziomu.

Przykład 9

IF a(v99) THEN 'Lack of date' FI

Tekst "Lack of date" będzie wyświetlony tylko wtedy, jeśli pole 99 będzie nieobecne w formatowanym rekordzie.

Przykład 10

if v11:'a' then 'Analytical' fi, if v11:'m' then 'Monograph' fi, if v11:'s' then 'Serial' fi, if v11:'c' then 'Collection' fi

Dla podanego pola (v11) z zakodowanym opisem bibliograficznym zostanie wyświetlony tekst

"Analytical" jeśli podany jest kod a

"Monograph" jeśli podany jest kod M

"Serial" jeśli podany jest kod S

"Collection" jeśli podany jest kod "C".

Przykład 11

```
if p(V10^b) and p(v10^d) then V10^a" // ",V10^c else V10^a" – ",V10^c  
fi
```

Zawartość podpól a oraz c pola o wyróżniku 10 będzie oddzielona podwójnym ukośnikiem '/' tylko jeśli dwa podpola $\wedge b$ oraz $\wedge d$ są obecne w rekordzie, jeśli przynajmniej jedno z podpól jest nieobecne ($\wedge b$ lub $\wedge d$) podpola $\wedge b$ oraz $\wedge d$ będą oddzielone znakiem '-'

Język wyszukiwawczy CDS/ISIS

Aby dokonywać zaawansowanego wyszukiwania w bazie CDS/ISIS należy sformułować zapytanie w języku wyszukiwawczym CDS/ISIS. Język wyszukiwawczy wykorzystuje operatory Boolowskie, dzięki czemu istnieje możliwość wygodnego zapisu operacji logicznych dotyczących klas obiektów. Każdy termin wyszukiwawczy związany z danym rekordem może być w rzeczywistości postrzegany jako reprezentant klasy wszystkich rekordów związanych z danym terminem. Tak więc przez wskazanie operacji logicznych między terminami wyszukiwawczymi można precyzyjnie określić klasę rekordów, które mają być znalezione w procesie wyszukiwawczym.

W niniejszym aneksie nie zamierzamy przedstawiać szczegółowego sposobu formułowania strategii wyszukiwawczych, lecz jedynie zilustrować możliwości języka wyszukiwawczego CDS/ISIS. Oczywiście sama znajomość języka wyszukiwawczego jest niewystarczająca, aby efektywnie wyszukiwać w bazie. Konieczna jest również znajomość terminów wyszukiwawczych, a także reguł indeksowania rekordów w danej bazie.

Najbardziej efektywnym sposobem przeprowadzenia wyszukiwania jest wykorzystanie zbioru odwróconego (jednym z powodów utrzymywania zbioru odwróconego jest szybkie wyszukiwanie). Jednak proces ten jest możliwy pod warunkiem, gdy zapytania wyszukiwawcze użytkownika mogą być wyrażone przez terminy zbioru odwróconego, a to oznacza, że wymagania te winny być już przewidziane przez projektanta bazy. Jednak bez względu na to, jak dobrze została pomyślana struktura bazy, istnieje prawdopodobieństwo, że użytkownik nie będzie mógł sformułować zapytania przy wykorzystaniu terminów ze zbioru odwróconego. W takich przypadkach CDS/ISIS daje możliwość wykorzystania alternatywnej metody zwanej wyszukiwaniem pełnotekstowym, która umożliwia formułowanie dowolnych zapytań niezależnie od zawartości zbioru odwróconego. Jednak metoda ta jest mało efektywna i zwykle jest wykorzystywana wraz z wyszukiwaniem za pomocą zbioru odwróconego jako część strategii wyszukiwawczej (choć może być używana niezależnie). Jeśli okaże się, że w nowo stworzonej bazie wyszukiwanie pełnotekstowe wykonywane jest zbyt często, może to oznaczać konieczność przedefiniowania zbioru odwróconego.

Typy terminów wyszukiwawczych

Dostępne są trzy typy terminów wyszukiwawczych:

- terminy jednoznaczne;
- terminy jednoznaczne z prawym obciążeniem;
- terminy ANY.

Terminy jednoznaczne

Termin jednoznaczny jest dowolnym wyrażeniem wyszukiwawczym zdefiniowanym dla danej bazy. W momencie formułowania zapytania (przy wykorzystaniu terminów jednoznacznych) należy mieć pod ręką słownik terminów wyszukiwawczych (ze zbioru odwróconego). Słownik terminów ze zbioru odwróconego dostępny jest z pozycji menu (np. w wersji angielskiej "T – Display terms dictionary" z podsystemu ISISRET).

Należy pamiętać, że nawet najmniejsze różnice w pisowni nie są tolerowane przez system. Jeśli zatem w słowniku dany termin jest poprzedzony prefiksem, to również w zapytaniu konieczne jest zastosowanie tego prefiksu.

Jeśli termin wyszukiwaczy zawiera w sobie nawiasy lub jakieś operatory wyszukiwawcze (* + (G) (F) . \$ ^) lub zaczyna się numerem oraz znakiem #, to należy znaki te umieszczać w podwójnych cudzysłowach. Tak więc jeśli termin wyszukiwaczy jest następujący:

GERMANY (FEDERAL REPUBLIC)

to w wyszukiwaniu należy zapisać:

"GERMANY (FEDERAL REPUBLIC)"

w przeciwnym przypadku CDS/ISIS wygeneruje komunikat o błędzie.

Terminy jednoznaczne z prawym obcięciem

Zamiast określenia terminu jednoznacznego możliwe jest podanie tylko jego początku. Technika ta określana jako obcinanie prawostronne, pozwala na przeszukiwania tylko początkowej sekwencji znaków. Prawostronne obcinanie jest uzyskiwane przez umieszczenie znaku dolara '\$' bezpośrednio za ostatnim znakiem początku terminu.

Przykład 12

Załóżmy, że słownik (zbiór odwrócony) zawiera m.in. następujące terminy:

FORESHORE
FOREST
FORESTER
FOREST-FLY
FORESTRY
FORETOP

Zatem zapisanie FOREST\$ jest równoważne:

FOREST
FORESTER
FOREST-FLY
FORESTRY

Terminy ANY

Termin ANY jest terminem zbiorczym odnoszącym się do wcześniej zdefiniowanych terminów wyszukiwawczych. Za każdym razem, kiedy używamy terminu ANY w wyrażeniu wyszukiwawczym, CDS/ISIS automatycznie dokonuje operacji sumy logicznej dla wszystkich terminów związanych z danym terminem ANY.

Termin ANY składa się ze słowa ANY, po którym następuje niepowtarzalny identyfikator zwykle mnemotechniczny przypisany odpowiedniemu zestawowi terminów. Na przykład termin KRAJE BENELUKSU może być wykorzystany przy wyszukiwaniu rekordów zaindeksowanych za pomocą nazw krajów tej grupy (Belgia, Holandia, Luksemburg).

Zanim będziemy mogli użyć terminu ANY, jego znaczenie musi być zdefiniowane przez określenie powiązanego z nim zbioru terminów. Oczywiście nie wszystkie bazy muszą implementować tę opcję. Zatem przed wyszukiwaniem z zastosowaniem terminu ANY należy upewnić się, czy jest ona zaimplementowana.

Operatory wyszukiwawcze

W wyrażeniu wyszukiwawczym można łączyć dwa lub więcej terminów wyszukiwawczych używając operatorów wyszukiwawczych, które wskazują, jaki związek zachodzi między terminami. Istnieją trzy typy operatorów wyszukiwawczych:

- Dwa typy operatorów położenia.
- Operatory Boolowskie:

- Suma logiczna OR. Zapisywana przez znak plus '+'
- Iloczyn logiczny AND. Zapisywany przez znak '*'
- Dopełnienie logiczne NOT. Zapisywane przez znak '^'

Operatory położenia dla pól

(G) – termin połączony tym operatorem musi pojawiać się w tym samym polu. Na przykład wyrażenie: woda (G) ziemia – wyszuka wszystkie rekordy, które zawierają zarówno termin 'woda' jak i 'ziemia' w tym samym polu.

(F) – termin połączony tym operatorem musi pojawiać się w tym samym wystąpieniu powtarzalnego pola.

Na przykład wyrażenie: woda (F) ziemia – wyszuka wszystkie rekordy, które zawierają zarówno termin 'woda' oraz 'ziemia' w tym samym polu lub w tym samym wystąpieniu powtarzalnego pola.

Oba operatory zastosowane do pól niepowtarzalnych będą dawały taki sam efekt.

Operatory położenia

"." – (kropka) podobnie jak (F), z dodatkowym ograniczeniem: słowa nie powinny być od siebie oddalone o więcej niż n słów, gdzie n jest liczbą kropek między słowami minus 1. Np.

A . B słowa występują obok siebie

A . . B – a i B są oddzielone od siebie najwyżej jednym słowem

A . . . B a i B są oddzielone od siebie najwyżej dwoma słowami

"\$" – (znak dolara) podobnie jak (F), z dodatkowym ograniczeniem: słowa powinny być oddalone od siebie dokładnie o n słów, gdzie n jest liczbą kropek między słowami minus 1.

A \$ B słowa występują obok siebie

A \$ \$ B – A i B są oddzielone od siebie dokładnie jednym słowem,

A \$ \$ \$ B A i B są oddzielone od siebie dokładnie dwoma słowami.



UWAGA:

Należy pamiętać, że przed oraz po każdym z operatorów '.' oraz '\$' musi występować jedna spacja.

Składnia wyrażeń wyszukiwawczych

W trakcie obliczania logicznej wartości wyrażeń zachowany jest następujący priorytet:

Priorytet	Operator
(najwyższy)	\$ oraz .
↓	(F)
↓	(G)
↓	* oraz ^
(najniższy)	+

W bardziej zaawansowanych zapytaniach korzysta się z nawiasów, które można zagnieżdżać. Należy jednak pamiętać, aby liczba otwartych nawiasów (lewostronnych) była taka sama, jak liczba zamkniętych nawiasów (prawostronnych).

Jeśli w danym wyrażeniu pojawiają się obok siebie więcej niż raz operatory wyszukiwawcze o tym samym priorytecie (w obrębie jednych nawiasów), to wyrażenia te będą wykonywane w kolejności: od strony lewej do prawej. Załóżmy, że zbiór odwrócony zawiera następujący zbiór terminów:

D=ROLNICTWO
D=WĘGRY
D=RYNEK PRACY
D=POLSKA

Aby uzyskać informacje na temat rynku pracy, to winniśmy sformułować następujące zapytanie:

D=RYNEK PRACY

Aby uzyskać informacje na temat dwóch krajów: Polski i Węgier, należy zapisać:

D=WĘGRY+ D=POLSKA

Aby odnaleźć informacje na temat rolnictwa Węgier, należy zapisać:

D=ROLNICTWO* D=WĘGRY

Aby odnaleźć informacje na temat rolnictwa dwóch krajów: Węgier oraz rolnictwa Polski, należy zapisać:

(D=WĘGRY+ D=POLSKA) * D=ROLNICTWO

Jeśli pominiemy nawiasy w ostatnim przykładzie, rezultat wyszukiwania będzie błędny: uzyskamy informacje na temat Węgier (nie tylko rolnictwa) oraz rolnictwa Polski.

Strategia wyszukiwawcza

Do każdego zapytania CDS/ISIS przyporządkowuje niepowtarzalny numer i zapamiętuje wyniki zapytań. W późniejszych zapytaniach wyszukiwawczych użytkownik może odwołać się do wcześniej wprowadzonych zapytań, używając niepowtarzalnego numeru zapytania poprzedzonego znakiem '#'.

Możliwość ta pozwala na tworzenie strategii wyszukiwawczej krok po kroku (tj. tworzenia najpierw prostych zapytań oraz ich łączenia w bardziej złożone zapytania wykorzystując jedynie ich numery).

Przyjmijmy, że zamierzamy znaleźć informacje na temat rolnictwa w Polsce oraz na Węgrzech. Mamy dwie możliwości:

1) możemy sformułować złożone zapytanie:

(D=WĘGRY+ D=POLSKA) * D=ROLNICTWO

2) lub sformułować szereg prostszych zapytań

(#1) D=WĘGRY
(#2) D=POLSKA
(#3) D=ROLNICTWO
(#4) #1 + #2

Rezultat ostatniego zapytania będzie taki sam jak w 1):

#3 * #4

Wyszukiwanie pełnotekstowe

Wyszukiwanie pełnotekstowe pozwala na przeszukiwanie pól, których zawartość nie została wprowadzona do zbioru odwróconego oraz na określenie warunków, niemożliwych do ustalenia przez wyrażenia wyszukiwawcze w połączeniu z terminami ze zbioru odwróconego, jak np. porównanie wartości numerycznych z danych pól.

Należy jednak pamiętać, że ponieważ wyszukiwanie pełnotekstowe sprawdza całą bazę w sposób sekwencyjny (rekord po rekordzie), taki proces wyszukiwawczy może być niezwykle czasochłonny i zależy w dużej mierze od ilości rekordów.

Ogólna forma wyszukiwania pełnotekstowego wygląda następująco:

? [#n] wyrażenie wyszukiwawcze

? – jest wskaźnikiem wyszukiwania pełnotekstowego;

[#n] jest opcjonalnym parametrem umożliwiającym zawężenie zakresu przeglądanych rekordów do wyników uprzednio wprowadzonego pytania (n jest numerem zapytania, odnoszącym się do zbioru odwróconego lub numerem zapytania pełnotekstowego). Jeśli składnik ten jest pominięty przeszukiwana jest cała baza. Wyrażenie wyszukiwawcze może być łączone operatorami Boolowskim.

Wyrażenia mogą przyjmować następującą postać:

vt : 'ciąg znaków' (pole zawierające znacznik t zawiera określony ciąg znaków);

p(vt) (pole zawierające znacznik t występuje w rekordzie) ;

a(vt) (pole zawierające znacznik t nie występuje w rekordzie) ;

val(vt) > wartość numeryczna.



UWAGA:

Wyrażenia Boolowskie mogą być dość złożone. Zamiast +/*/^ należy używać operatorów: **AND**, **OR** oraz **NOT**

Wyszukiwania przy wykorzystaniu zbioru odwróconego (standardowe) oraz pełnotekstowe nie powinny występować w tym samym wyrażeniu wyszukiwawczym.

Przykład 13

Aby znaleźć wszystkie rekordy zawierające kod 'A' w polu 11, należy użyć następującego zapytania:

? 11: 'A'

Przykład 14

Aby znaleźć te rekordy, które w polu 11 nie są puste, należy użyć następującego zapytania:

? p(v11)

Przykład 15

Aby znaleźć te rekordy, które w polu 11 nie są puste, zaś pole 10 nie występuje, należy użyć następującego zapytania:

p(v11) and a(v10)

Przykład 16

Aby znaleźć te rekordy, zawierające wartość numeryczną w polu 1403 większą niż 1988, należy użyć następującego zapytania:

```
? val(v1403) > 1988
```

Przykład 17

Aby stwierdzić, że np. pole o znaczniku v11 nie zawiera tekstu 'rolnictwo', należy zapisać:

```
? p(V11) and not (V11:'rolnictwo')
```

Przykład 18

Aby znaleźć rekordy zawierające tekst fishery przynajmniej w jednym z wymienionych pól: 10, 15, 20

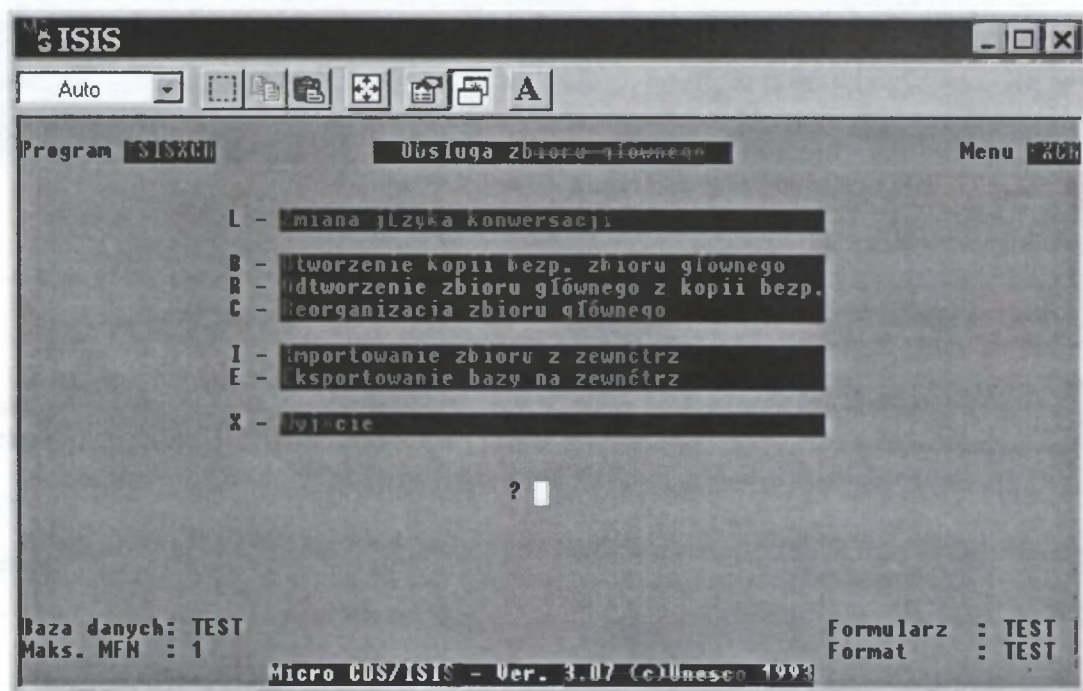
```
? S(v10, v15, v20): 'nawozy'
```

Zapis ten jest równoważny zapisowi poniżej:

```
? v10:'nawozy' or v15:'nawozy' or v20:'nawozy'
```

Eksport, import danych

Istotną kwestią w systemie WWW-ISIS jest eksport oraz import danych. Dane można wyeksportować do dowolnego formatu przy wykorzystaniu poleceń wydruku oraz języka formatowania. Jednak w celu przenoszenia danych między różnymi bazami WWW-ISIS konieczne jest wykorzystanie funkcji M (obsługa zbioru głównego, rys. 80) pozwalającej na dokonywanie m.in. eksportu i importu danych.



Rys. 80. Obsługa zbioru głównego

Eksport danych pozwala na zapisanie danych w zbiorze zgodnym z formatem ISO2709. Dzięki dodatkowym opcjom m.in. „zbiór parametrów formatowania możliwe jest wyeksportowanie dowolnego podzbioru zbioru głównego przez podanie nazwy uprzednio zdefiniowanego pliku fst.

ISIS.exe

Wymiana danych - Parametry eksportu

Nazwa bazy danych **ESI**

Parametry zbioru w formacie ISO

Nazwa zbioru wyjściowego **E1.ISO** Separator pola
Separator rekordu

Parametry wyboru rekordów

Zakres MFN **1/32000** Zbiór wyników Zbiór kluczy <Y/N>

Parametry reformatowania

Zbiór parametrów reformatowania **transl.fst** Tablica konwersji
Numer pierwszego rekordu

Rys. 81. Ustalenie parametrów eksportu

W przypadku systemu WWW-ISIS z opcji tej korzystamy m.in. w momencie implementacji tezauryśa stworzonego za pomocą programu MTM4. W takim przypadku konieczne jest wyeksportowanie w programie MTM4³ tezauryśa do postaci zgodnej z wersją MTM3, co wymaga podania w zbiorze parametrów formatowania nazwy zbioru reformatującego FST o nazwie transl.fst, a następnie wczytania tak wyeksportowanego zbioru do odpowiedniego systemu ISIS przez opcję importuj (rys. 81).

ISIS

Wymiana danych - Parametry importu

Nazwa bazy danych **ESI**

Parametry zbioru w formacie ISO

Nazwa zbioru wejściowego w formacie ISO **E1.ISO** Separator pola
Separator rekordu

Parametry ładowania

Numer pierwszego rekordu lub wyróżnik pola zawierającego MFN
[E] Ładuj / [D] Bełćcz / [W] Wymieł

Parametry reformatowania

Zbiór parametrów reformatowania Tablica konwersji

ast(owanie znaków Ostatnia strona

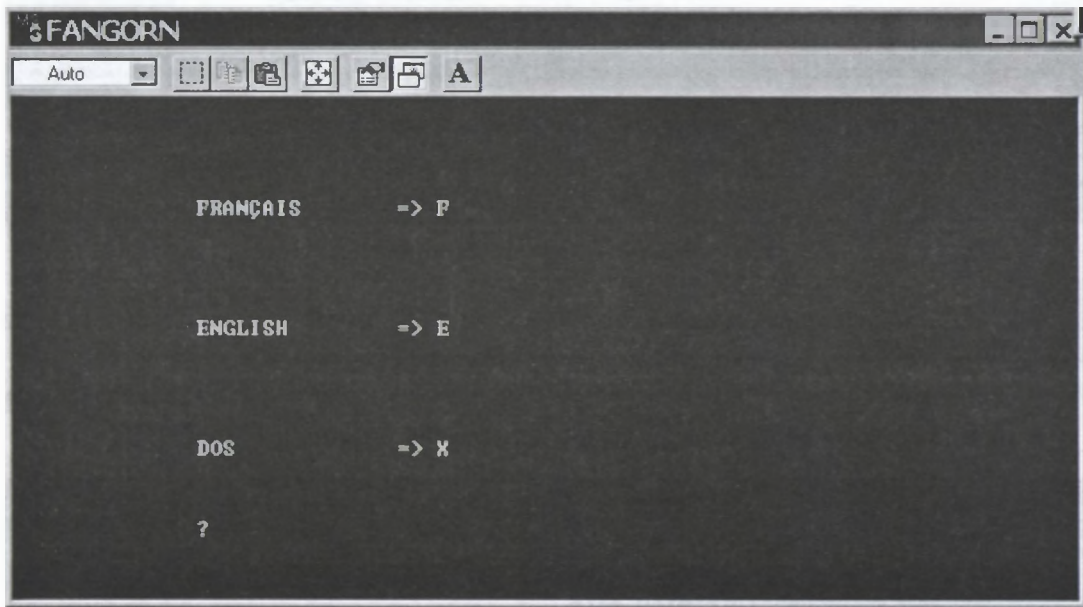
Rys. 82. Ustalenie parametrów importu

³ MTM4 jest aplikacją ISIS, tak więc część opcji programu jest zgodna ze standardem ISIS.

Aneks C

Konwersja danych do formatu ISO2709

Poniżej zostanie przedstawiona przykładowa metoda konwersji danych do formatu zgodnego ze standardem ISO2709 akceptowanym przez CDS/ISIS. W tym celu wykorzystamy ogólnodostępny program FANGORN.



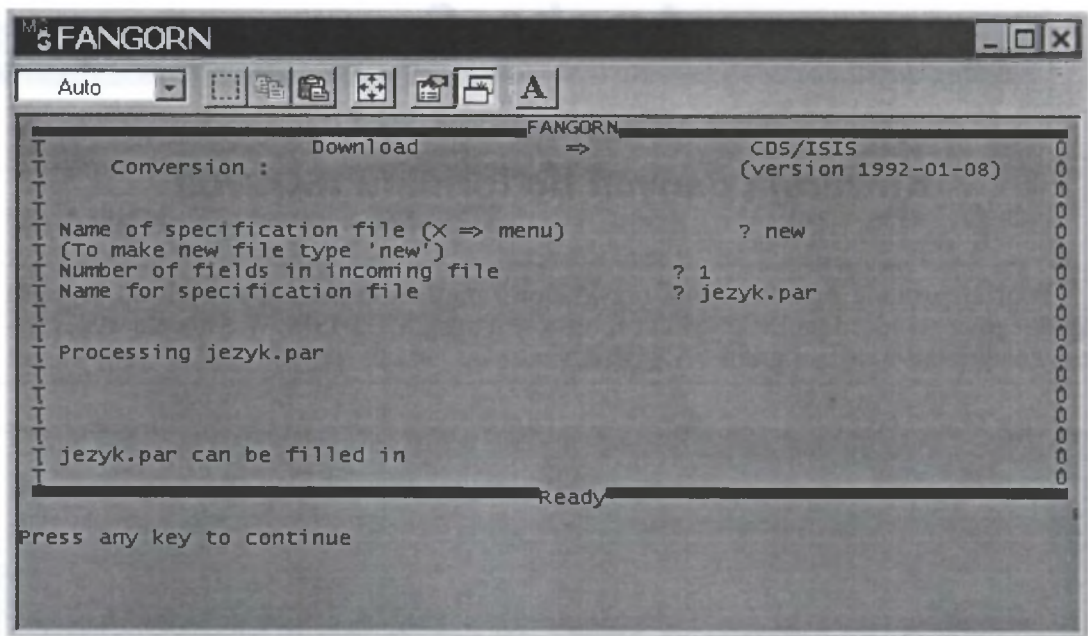
Rys. 83. Program FANGORN – ustalenie języka konwersacji

Krok pierwszy – ustalenie parametrów konfiguracyjnych

Zanim przystąpimy do stworzenia odpowiedniego zbioru z danymi, niezbędne jest ustalenie odpowiednich parametrów, potrzebnych do dokonania konwersji, jak: znacznik rekordu, znaczniki pól (podpól) itd.

Założmy, że naszą przykładową bazą jest baza języków (z bazy Cordis UE). Jest to prosta baza z jednym atrybutem – „nazwą języka”.

Zanim program dokona konwersji bazy należy określić zbiór z parametrami. Jest to proces częściowo zautomatyzowany.



Rys. 84. FANGORN – Tworzenie zbioru konfiguracyjnego

W tym momencie mamy już utworzony zbiór JEZYK, którego parametry należy teraz szczegółowo ustalić.

Tabela 21

Przykładowy zbiór konfiguracyjny lang.par dla programu FANGORN

Komentarz	Przykładowy zbiór konfiguracyjny lang.par dla programu FANGORN	
	Parametry zbioru	Przykładowe Wartości
	SPECIFICATION FILE	
	Which conversion is specified	:
Liczba pól	Number of fields	: 1
Znacznik początku rekordu	Tag indicating start/end of record	: Record -
	Line indicating start/end of record	:
	Two texts indicating start/end of rec.:	
Znacznik w zbiorze poddawanych konwersji	Tag in incoming file	: LANG:
Znacznik w zbiorze wyjściowym	Tag in ISO-2709 file	: 10
	(The following is not mandatory)	
	Spaces in continuation lines	:
	Subfield delimiter in incoming file	:
	Subfield delimiter ISO-2709 file	:
	Delimiter occurrences incoming file	:
	Texts to replace	:

Poniżej omówione są najistotniejsze parametry (por. tab. 21):

- **Liczba pól** – określa ogólną liczbę pól poddawanych konwersji.
- **Znacznik początku rekordu** – określa początek każdego rekordu.

- **Znacznik w zbiorze poddawany konwersji** – określa znacznik dla każdego pola w zbiorze wejściowym.
- **Znacznik w zbiorze wyjściowym** – określa znacznik dla każdego pola w zbiorze wynikowym ISO2709.

Przytoczony przykład jest jednym z wielu i nie musi stanowić wzorcowego punktu odniesienia. W przypadku ustalania parametrów konfiguracyjnych należy brać pod uwagę przede wszystkim: liczbę pól, znacznik początku rekordu, znacznik pola w zbiorze wejściowym oraz w zbiorze wynikowym, a w przypadku podpól znacznik podpola w zbiorze wejściowym oraz w zbiorze wynikowym. Zanim przystąpimy do konwersji warto jest przeanalizować dane w zbiorze wejściowym i sprawdzić, czy nie istnieją w nim elementy będące „kandydatami” na odpowiednie znaczniki w zbiorze konfiguracyjnym.

Poniżej znajduje się fragment z bazy Dialog⁴.

```

9/4/1 of 550
EN- EJ375296|
AN- IR518596|
TI- a Microcomputer-Based Interactive Presentation Development System.|
AU- Moreau, Dennis R.; Dominick, Wayne D.|
9/4/2 of 550
EN- EJ374152|
AN- SE542794|
TI- Software Reviews.|
AU- Smith, Richard L., Ed.|

```

W podanym przykładzie z Dialog znacznik początku rekordu możemy określić następująco:

: .00/00~~/4/

Krok drugi – umiejscowienie danych w tabeli

W przypadku konwersji danych z bazy Microsoft Access lub innej relacyjnej bazy danych np. Excell, zaleca się stworzenie dwuwymiarowej tabeli danych. W tabeli 22 podany jest przykład danych z bazy Cordis (Unii Europejskiej), które zostały wprowadzone do jednowymiarowej tablicy w bazie Excell.

Tabela 22

Przykładowe dane w arkuszu Excella

Dane do konwersji FANGORN – znacznik początku rekordu	Dane do konwersji FANGORN znacznik Pola	Dane z bazy Cordis
1	2	3
Record -	LANG:	Abkhazian
Record -	LANG:	Afar
Record -	LANG:	Afrikaans
Record -	LANG:	Albanian
Record -	LANG:	Amharic

⁴ Przykład zaczerpnięty z dołączanej do programu FANGORN dokumentacji (Besemer H., Nieuwenhuysen P. : How to use Fangorn – a microcomputer program, for the conversion of databases, to ISO-2709 format).

1	2	3
Record -	LANG:	Arabic
Record -	LANG:	Armenian
Record -	LANG:	Assamese
Record -	LANG:	Avestan
Record -	LANG:	Aymara
Record -	LANG:	Azerbaijani
Record -	LANG:	Bashkir
Record -	LANG:	Basque
Record -	LANG:	Belarusian
Record -	LANG:	Bengali
Record -	LANG:	Bihari
Record -	LANG:	Bislama
Record -	LANG:	Bosnian
Record -	LANG:	Breton
Record -	LANG:	Bulgarian
Record -	LANG:	Burmese
Record -	LANG:	Catalan
Record -	LANG:	Chamorro
Record -	LANG:	Chechen
Record -	LANG:	Chichewa
Record -	LANG:	(...)

Krok trzeci – zapisanie danych w zbiorze tekstowym

W przypadku konwersji danych z bazy Microsoft Access lub innej relacyjnej bazy danych np. Excell, należy umieścić dane w zbiorze tekstowym, którego przykładowy fragment zamieszczamy poniżej.

```
Record -
LANG: Abkhazian
Record -
LANG: Afar
Record -
LANG: Afrikaans
Record -
LANG: Albanian
Record -
LANG: Amharic
Record -
LANG: Arabic
Record -
LANG: Armenian
(...)
```

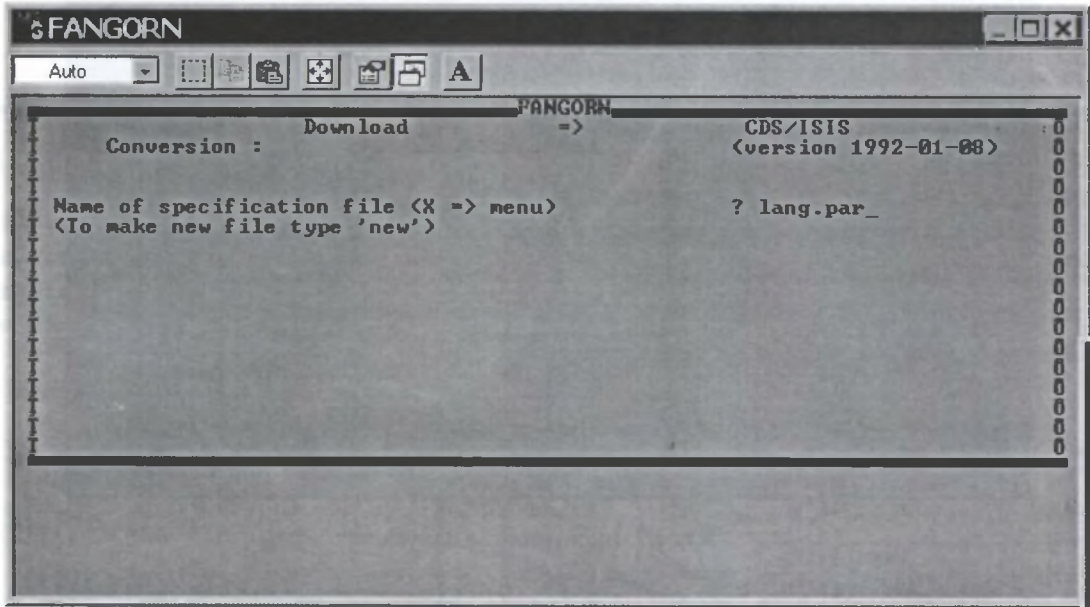
Krok czwarty – przeprowadzenie konwersji

W ostatnim etapie przeprowadzamy konwersję danych zgodnie z parametrami w zdefiniowanym wcześniej zbiorze konfiguracyjnym. Uruchamiamy program FANGORN, po czym następuje standardowo zapytanie o język konwersacji (patrz rys. 83). Można też uruchomić program z parametrem określającym zbiór konfiguracyjny oraz język konwer-

sacji np. przy założeniu, że naszym zbiorem konfiguracyjnym był zdefiniowany w pierwszym kroku 1, a językiem konwersacji jest angielski, winniśmy napisać w wierszu poleceń (w katalogu programem FANGORN oraz ze wszystkimi zbiorami do konwersji lub z podaniem ścieżki dostępu do programu FANGORN):

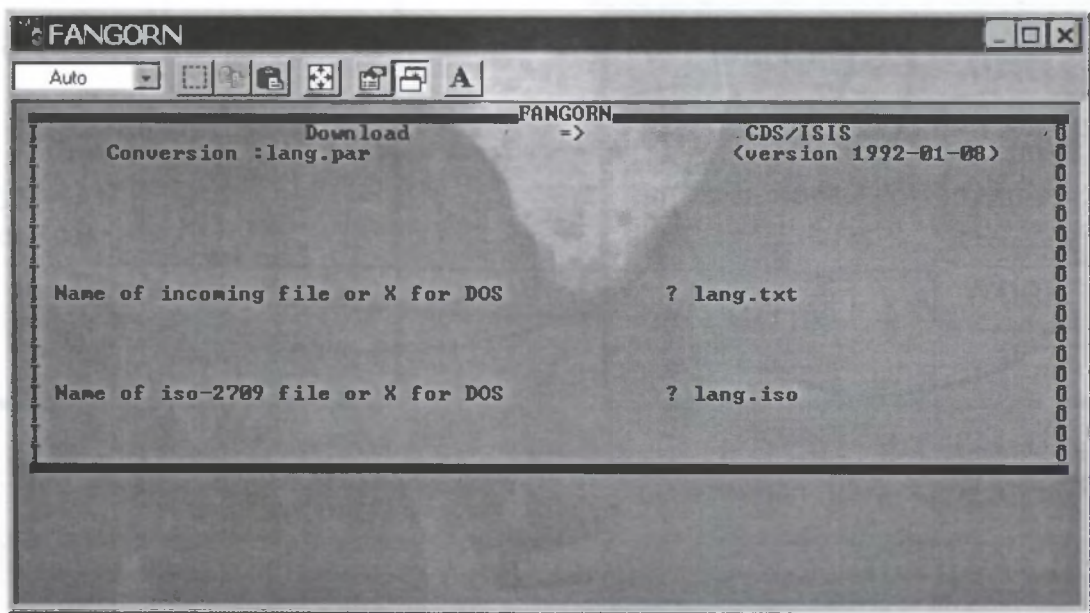
```
fangorn l:e s:lang.par
```

Jeśli uruchomimy program FANGORN bez żadnego parametrów, po ustaleniu języka konwersacji (patrz rys. 85) nastąpi uruchomienie programu z prośbą o podanie nazwy zbioru konfiguracyjnego (podajemy zgodnie z przykładem lang.par).



Rys. 85. Podanie nazwy zbioru konfiguracyjnego

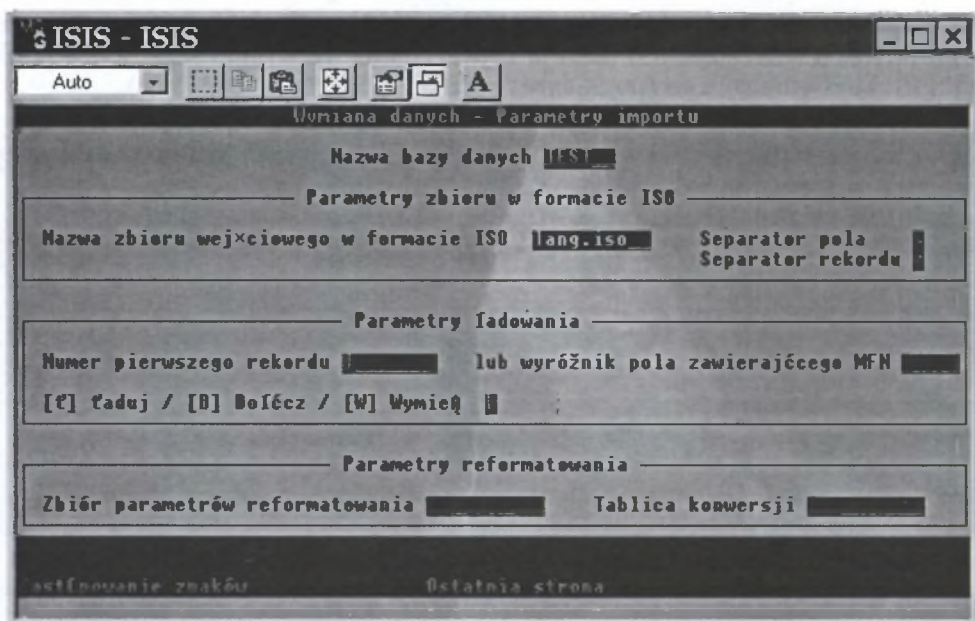
Następnie program żąda podania nazwy zbioru wejściowego (Name of incoming file or X for DOS...) oraz nazwy zbioru wyjściowego (Name of iso-2709 file or X for DOS..) (rys. 86).



Rys. 86. Przeprowadzenie konwersji

Import danych

Po wygenerowaniu danych w formacie ISO 2709 można zaimportować dane do bazy. Należy jednak pamiętać, aby nasz zbiór w formacie ISO znajdował się we właściwym katalogu⁵.



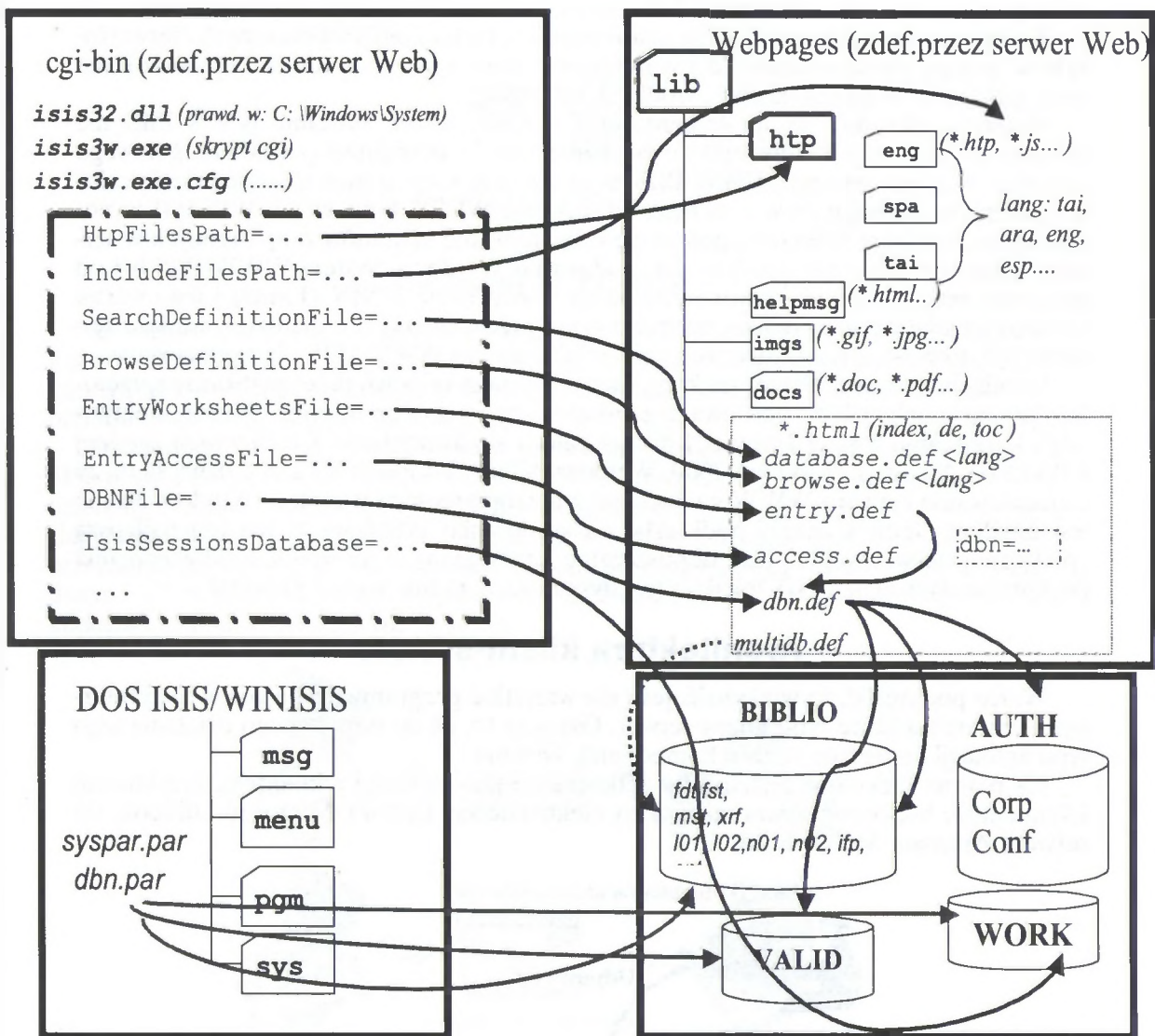
Rys. 87. Importowanie danych

⁵ Przeważnie jest to katalog, w którym znajduje się zbiór isis.exe.

Aneks D

Zależności konfiguracyjne

Zależności między wybranymi zbiorami konfiguracyjnymi w WWW-ISIS



Aneks E

Środowisko sieci

Wybrane aspekty konfiguracyjne, serwery WWW

Zastosowanie technologii internetowych czyni dostęp do zasobów i ich utrzymanie znacznie łatwiejszymi i tańszymi głównie z dwóch powodów: upowszechnienie się technologii internetowych jako standardu *de facto*, a co za tym idzie dynamicznie rozwijającego się rynku produktów, w tym przebogatej oferty produktów darmowych (*freeware* oraz *open source* a także *shareware* lub *adware*) oraz ich coraz większej znajomości.

Te dwa czynniki przemawiają za zastosowaniem technologii internetowych (serwerów WWW, poczty elektronicznej itd.) w miejscach, które wymagają tanich i sprawdzonych rozwiązań, czyli w bibliotekach i ośrodkach informacji.

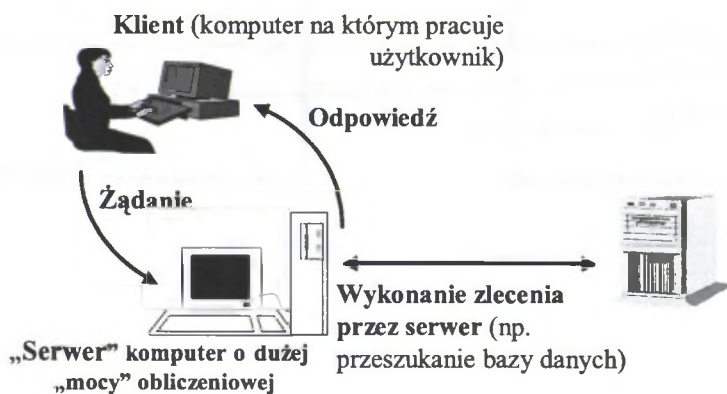
Należy podkreślić, że implementacja i przetestowanie systemu WWW-ISIS nie wymaga od użytkownika dostępu do sieci Internet. W momencie posiadania gotowego i przetestowanego systemu WWW-ISIS można zacząć rozpatrywać umiejscowienie całości systemu na działającym w sieci Internet serwerze WWW. Ponieważ WWW-ISIS wykorzystuje technologie Windows, poleca się rozpatrywanie systemów zaopatrzonych w serwery Microsoft IIS, czyli działające w środowisku Windows. System WWW-ISIS był od początku swego istnienia rozwijany wspólnie z platformą UNIX (Linux) i jego wersja Linuxowa nadal istnieje. Jednak podręcznik ten przeznaczony jest dla użytkowników systemu Windows, stąd też Linuxowe „wcielenia” systemu WWW-ISIS nie są omawiane.

Instalacja serwera WWW na komputerze wymaga odpowiedniej platformy sprzętowej, ponadto należy brać pod uwagę zainstalowany system operacyjny. W celu efektywnego korzystania z sieci Internet/Intranet zaleca się korzystanie z technologii zgodnej z Windows NT (czyli Windows 2000, Windows NTv.4, Windows XP). Nie oznacza to, że zainstalowanie serwera WWW na komputerze zaopatrzonym w system Windows 9x nie jest możliwe. Jednak należy podkreślić, że środowisko Windows 9x nie jest najlepszą „platformą serwerową” i jest dopuszczalne przeważnie w przypadku udostępniania (wykorzystywania) baz ISIS lokalnie (na dysku twardym lub przez CD-ROM).

Architektura klient-serwer

Warto podkreślić, że większość, jeśli nie wszystkie programy działające w sieci Internet mają architekturę typu klient-serwer. Oznacza to, że do poprawnego działania tego typu aplikacji konieczne są dwa komponenty systemu.

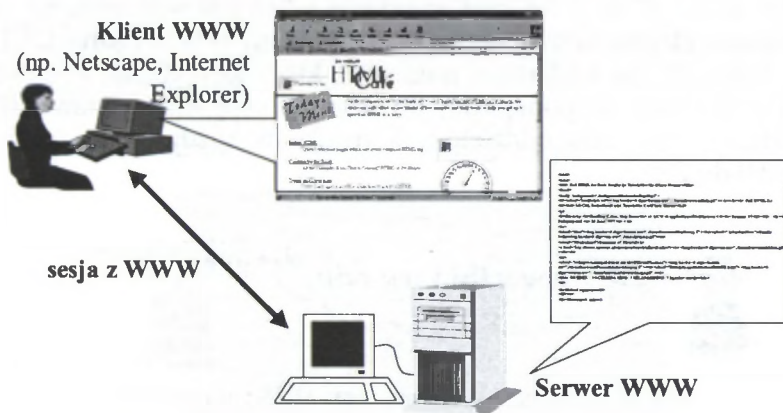
Na pewno większość czytelników zetknęła się przynajmniej z klientem. Przykładem klienta może być oprogramowanie poczty elektronicznej Eudora, Microsoft Outlook, zaś serwera Altavista Mail, IIS itd.



Rys. 88. Architektura klient-serwer

Zakładamy, że do poprawnego funkcjonowania poczty elektronicznej, większość czytelników musiała konfigurować swoje programy pocztowe tak, aby możliwe było wysyłanie i odbieranie np. poczty elektronicznej.

W przypadku sieci WWW klientem będzie znany każdemu Internet Explorer, Netscape, Opera itd. zaś serwerem IIS, Xitami, Netscape itd.

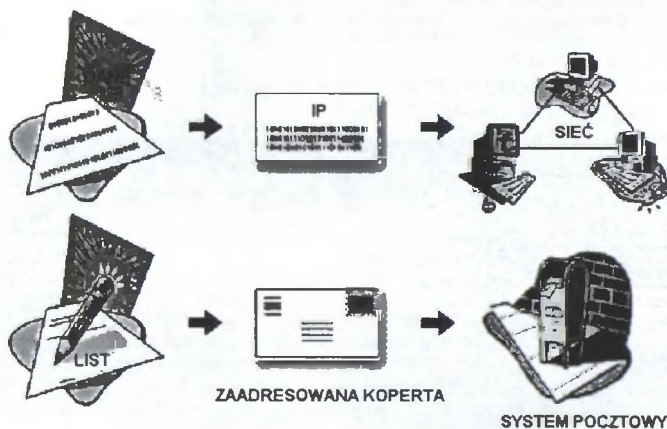


Rys. 89. Architektura klient-serwer, system WWW

Protokół TCP/IP

W aneksie tym nie zamierzamy wyczerpująco omówić całości zagadnień związanych z protokołem TCP/IP. Zostaną jedynie krótko zarysowane zagadnienia dotyczące takiej konfiguracji sieci, aby instalacja i zaimplementowanie systemu WWW-ISIS oraz jego przetestowanie było w pełni możliwe.

Protokół TCP/IP jest podstawowym protokołem komunikacyjnym w sieci Internet. Składa się z dwóch komponentów: protokołu TCP oraz IP, oba działają wspólnie w celu dzielenia danych na mniejsze porcje, które mogą być sprawnie przesłane przez sieć; określają punkt docelowy przesyłanych danych, sprawdzają czy dane dotarły do miejsca swego przeznaczenia; dokonują rekonstrukcji danych do oryginalnej postaci. Relacja między danymi, IP i siecią jest często porównywana do zależności między listem, zaadresowaną kopertą i systemem pocztowym (rys. 90).



Rys. 90. Zasada działania protokołu IP

Sam protokół IP daje następujące możliwości:

- służy do pakowania adresowania i danych;
- umożliwia sieci „czytanie” pakietów danych, przesyłanie ich dalej, do miejsca przeznaczenia;
- określa jak wiele danych może zawierać się w pojedynczym pakiecie.

Zaś protokół TCP:

- dzieli dane na pakiety, które mogą być efektywnie przesyłane przez sieć;
- sprawdza, czy wszystkie pakiety dotarły do miejsca swojego przeznaczenia;
- z otrzymanych pakietów na powrót „łączy” dane.

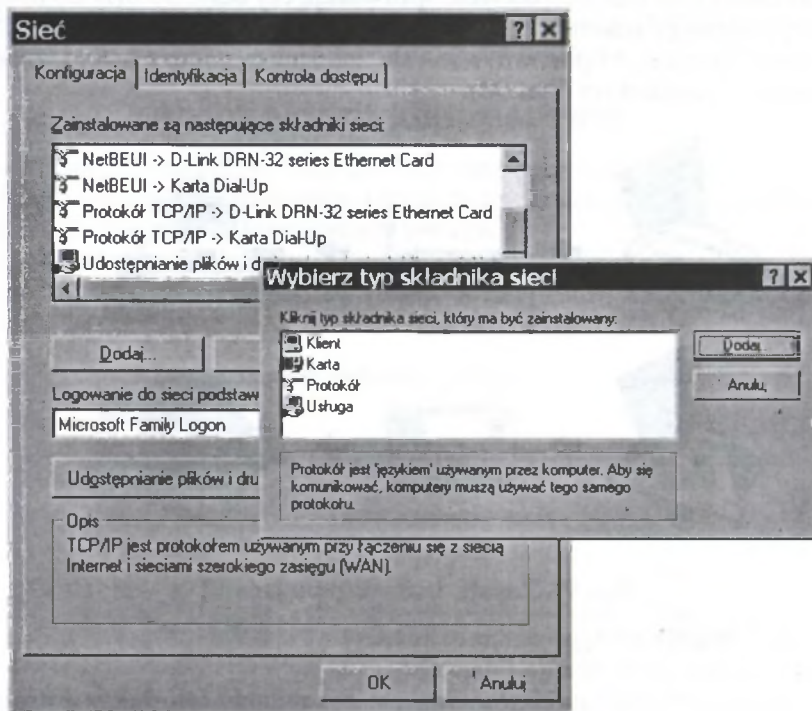
Aby pakiety mogły trafić do swego miejsca przeznaczenia, wszystkie komputery w sieci muszą posiadać adres. W sieci Internet wyrażenie adres jest wieloznaczne: może oznaczać adres w poczcie elektronicznej lub adres danej strony WWW (adres URL) lub adres serwera FTP. Adres IP jest unikalnym numerem, który identyfikuje komputery w sieci Internetem. Każdy komputer podłączony do sieci ma swój własny numer IP. Adres IP składa się z czterech numerów oddzielonych kropkami, każdy numer musi mieścić się w przedziale od 0 do 255.



Rys. 91. Przyporządkowanie nazw DNS i adresów IP

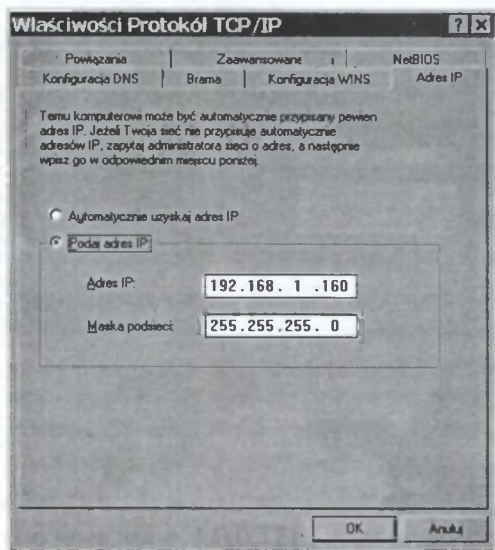
Aby ułatwić ludziom korzystanie z adresów internetowych wprowadzono tzw. system domen internetowych, w których każdy adres alfanumeryczny odpowiada adresowi IP (rys. 91). System domen (nazw internetowych) jest narzędziem sieciowym odpowiedzialnym za tłumaczenie wygodnych i w miarę zrozumiałych nazw komputerów w sieci (np. www.ala.org) na numeryczne adresy IP. Podobnie jak adresy IP tak i domeny internetowe są niepowtarzalne. Typowa domena ma charakter hierarchiczny: domena pierwszego rzędu i domeny kolejnych rzędów oraz nazwa komputera np.: www.ibin.uw.edu.pl. Typowa nazwa internetowa podaje typ organizacji lub kraj, w którym znajduje się dany komputer bazowy np.: (EDU, COM, GOV, UK, etc.).

W przypadku systemu Windows korzystanie z protokołu sieciowego TCP/IP wymaga jedynie zainstalowania odpowiednich komponentów systemu operacyjnego (rys. 92).



Rys. 92. Instalacja protokołu TCP/IP

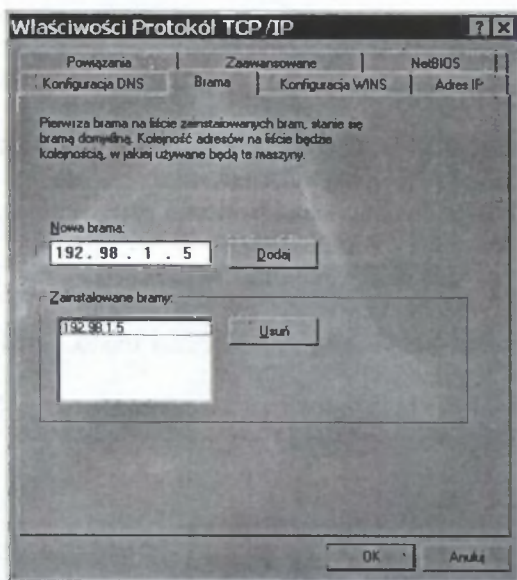
Konfiguracja protokołu TCP/IP nie powinna stanowić większego problemu. Przy założeniu, że nasz komputer nie będzie podłączony do sieci komputerowej i będzie jedynie pełnił rolę sprzętu do testowania aplikacji WWW-ISIS, wszelkie ustawienia mogą być wybrane w sposób dowolny, jednak z zachowaniem pewnych zasad. Do najważniejszych ustawień należy: adres IP, bramka DNS oraz adresy (IP) serwerów nazw domenowych. W przypadku instalacji protokołu TCP/IP na komputerze, który będzie pełnił rolę serwera internetowego adres IP powinien być niepowtarzalny i publiczny (widoczny w całej sieci Internet a nie tylko w podsieci) oraz wybrany zgodnie z ustaleniami administratora sieci, bądź bezpośrednio w porozumieniu z dostawcą usług internetowych. W tym samym oknie dialogowym ustawiana jest maska podsieci. Jeśli w danej podsieci ma być dostępna cała podsieć z puli 255 adresów, to wybierana maska wynosi 255.255.255.0. (rys. 93.)



Rys. 93. Konfiguracja maski podsieci i adresu IP

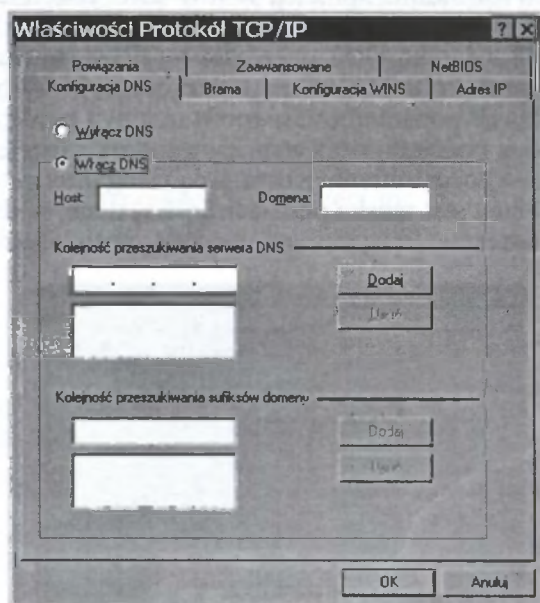
Warto pamiętać, że istnieje specjalny adres IP –127.0.0.1, który zawsze oznacza adres lokalnego komputera.

Kolejnym ustawieniem jest określenie IP bramy, czyli komputera, przez który będą przechodziły pakiety TCP/IP (rys. 94).



Rys. 94. Ustawienia bramy w protokole TCP/IP

W ostatnim etapie konfiguracji sieci należy ustalić adresy IP serwerów nazw domenowych: podstawowego oraz zapasowego (rys. 95).



Rys. 95. Określenie adresów IP serwerów DNS

W systemie Windows 9x istnieje specjalny zbiór „hosts” (przeważnie jest on nieaktywny i wówczas posiada rozszerzenie host.sam). Dzięki niemu odwołania do systemu nazw domenowych mogą być pomijane i ustalane na podstawie tego zbioru (np.: w zbiorze hosts zapisujemy w nowym wierszu „127.0.0.1 www.mojkomp.com”).

Serwery WWW

Poniżej zostaną omówione procesy konfiguracyjne 3 wybranych serwerów WWW: XITAMI, IIS oraz Sprite.

Serwer Xitami

Instalacja serwera XITAMI

Należy upewnić się, czy na komputerze nie jest zainstalowany żaden serwer, a następnie zainstalować serwer XITAMI przez uruchomienie instalatora. Instalacja jest dość prosta i sprowadza się praktycznie do uruchomienia programu instalacyjnego (znajdującego się na dołączonym krążku CD-ROM).

- Kliknąć podwójnie na ikonie programu instalacyjnego XITAMI bw3224d6.exe a następnie uważnie prześledzić procedurę instalacyjną.
- Wprowadzić i zapamiętać nazwę administratora i hasło (kiedy program instalacyjny tego zażąda).

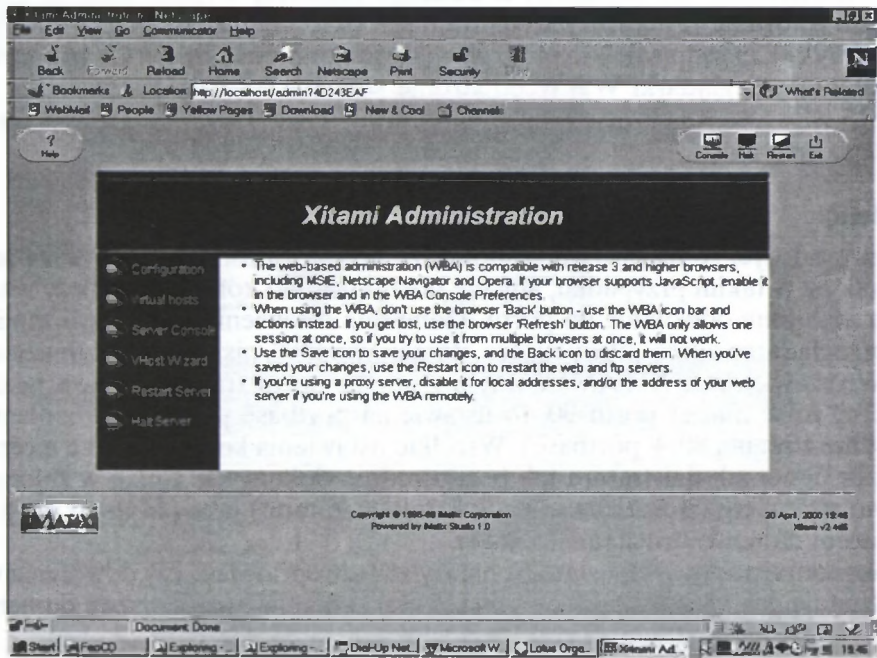
W przypadku ściągania instalatora z sieci należy połączyć się ze stroną www.xitami.com i stamtąd pobrać plik instalacyjny.

Ustawienia serwera

Zwykle serwer jest instalowany na dysku C w katalogu \Xitami, co jednak można zmienić w zależności od preferencji. Po zakończeniu procesu instalacyjnego i uruchomieniu serwera należy uruchomić przeglądarkę WWW (IE lub Netscape) i w polu loka-

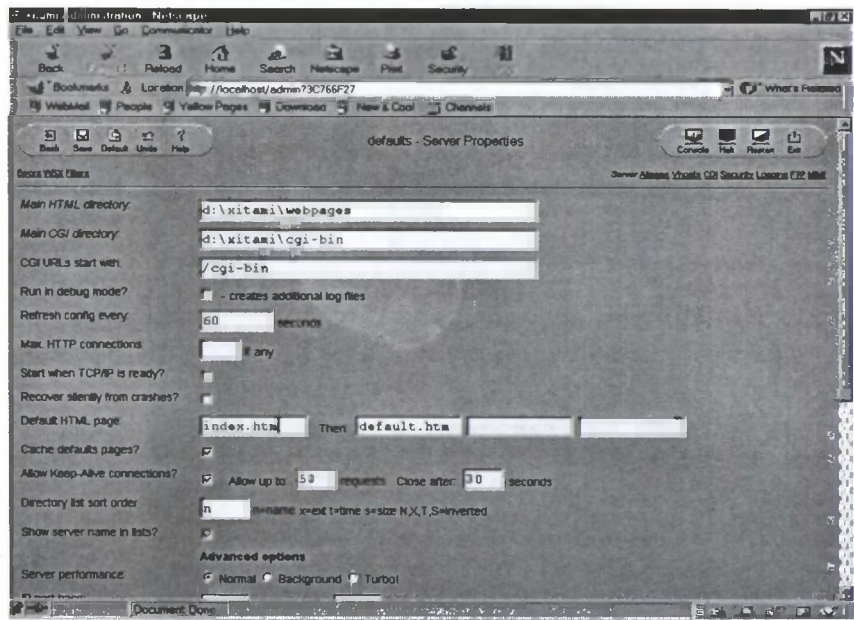
lizacji zasobów URL wpisać : `http://localhost`. Po chwili zgłosi się strona powitalna naszego nowo zainstalowanego serwera Xitami.

Strona powitalna zawiera większość istotnych informacji dotyczących samego serwera Xitami, najczęściej zadawane pytania, pomoc, obsługa i administracja serwera, warunki umowy itd. Ze strony tej możliwe jest przejście do strony administratora przez kliknięcie na "WEB-based admin", lub wpisanie adresu `http://localhost/admin`



Rys. 96. Powitalna strona administratora

Gdy próbujemy rejestrować się jako administrator serwera XITAMI, system zapyta nas o hasło, które zostało wprowadzone w czasie instalacji (w przypadku zapomnienia hasła można je odszukać w zbiorze konfiguracyjnym `defaults.aut`). Po wprowadzeniu poprawnej nazwy administratora oraz jego hasła system wyświetli ekran (rys. 96 i 97).



Rys. 97. XITAMI panel administracyjny – główne ustawienia serwera WWW

Po wybraniu opcji Configuration z panelu administracyjnego (rys. 96) ukaże się ekran z ważniejszymi ustawieniami konfiguracyjnymi serwera XITAMI (rys. 97). Większość z ustawień systemu powinna być zrozumiała dla przeciętnego użytkownika. Pierwszy parametr określa lokalizację katalogu głównego serwera WWW – skąd dokonywane są wszelkie odwołania do stron WWW i podkatalogów serwera WWW (zwykle jest to C:/Xitami/Webpages). Istotną rolę w strukturze katalogów odgrywać będzie katalog DEMO zlokalizowany (po instalacji) w katalogu głównym serwera wraz ze wszystkimi koniecznymi zbiorami. Kolejnym parametrem jest określenie katalogu cgi-bin – zawierającego skrypty cgi, czyli niewielkie programy (napisane przeważnie w C) wykonywane na żądanie klienta przeglądarki WWW po stronie serwera WWW. W przypadku systemu WWW-ISIS właściwa lokalizacja tego programu odgrywa decydującą rolę w funkcjonowaniu całego systemu.

Problemy

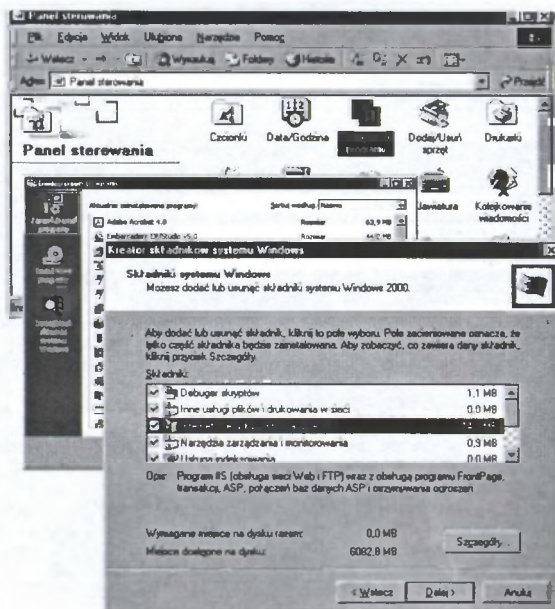
Możemy założyć, że na naszym komputerze będzie zainstalowanych kilka serwerów WWW. Jednak w takim przypadku, aby nie dochodziło do konfliktów należy ustawić dla każdego z nich inne odwołanie do portu. Domyślne ustawienia to 80, czyli zamiast wpisywać w przeglądarce adres „http://nazwa_domenowa:80/”, piszemy „http://nazwa_domenowa/”. W ustawieniach XITAMI domyślny numer portu traktowany jest jako 0, jeśli więc chcemy mieć numer portu 90, to ustawienia portbase powinniśmy zmienić na 10 (czyli zgodnie z regułą 80 + portbase). Wszelkie ustawienia konfiguracyjne można dokonywać przez panel administratora lub bezpośrednio dokonywać zmian w zbiorze konfiguracyjnym xitami.cfg (zlokalizowanym w katalogu Xitami) jako parametr portbase i za każdym razem na nowo uruchamiać serwer.

Po dokonaniu stosownych ustawień należy zawsze sprawdzić, czy odwołania do stron serwera działają poprawnie, zarówno lokalnie jak i zdalnie – przez nazwę domenową.

Serwer IIS (Internetowe Usługi Informacyjne)

Poniżej zostanie omówiona instalacja serwera IIS dla systemu operacyjnego Windows 2000 oraz wybrane aspekty jego funkcjonowania w systemie WWW-ISIS.

Instalacja

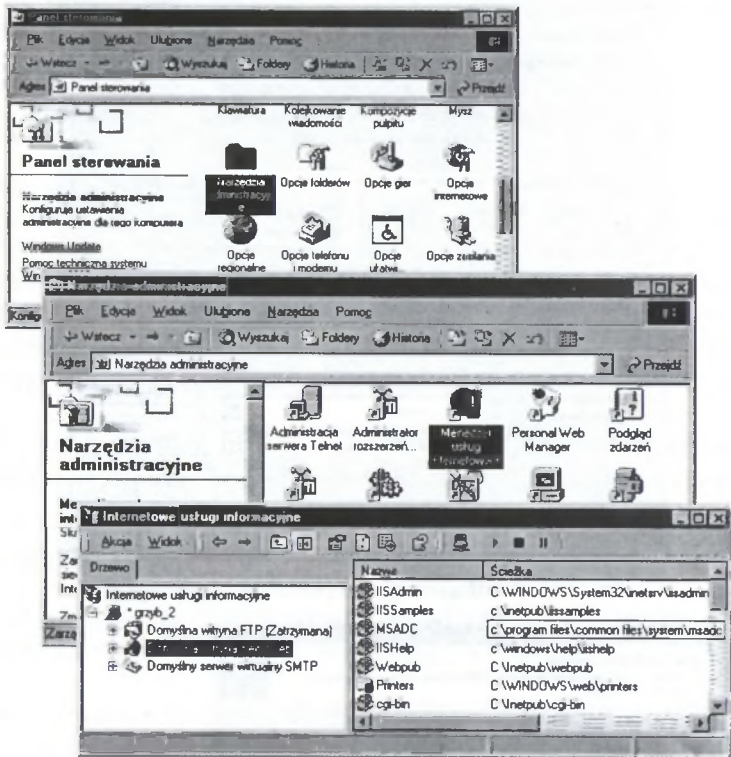


Rys. 98. Instalacja IIS (Internetowe Usługi Informacyjne)

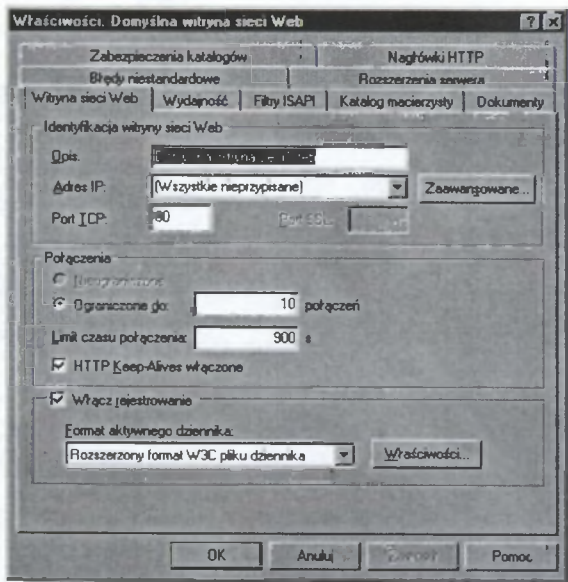
Instalacja serwera nie powinna stanowić większych problemów. Serwer IIS (Internetowe Usługi Informacyjne) jest poniekąd integralną częścią systemu Windows 2000. Jego instalacja dokonuje się z panelu sterowania (dodaj/usuń nowe programy, dodaj/usuń składniki systemu Windows rys. 98).

Konfiguracja

Dostęp do serwera IIS jest możliwy z poziomu panelu sterowania/narzędzia „administracyjne”.



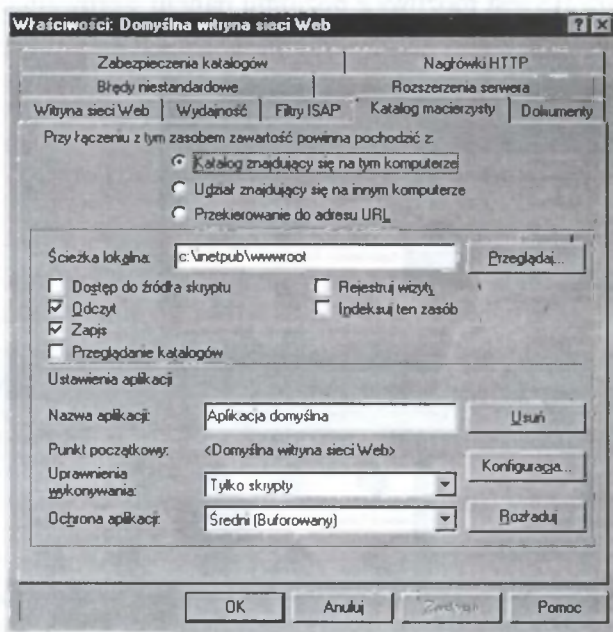
Rys. 99. Dostęp do IIS przez menedżera usług internetowych



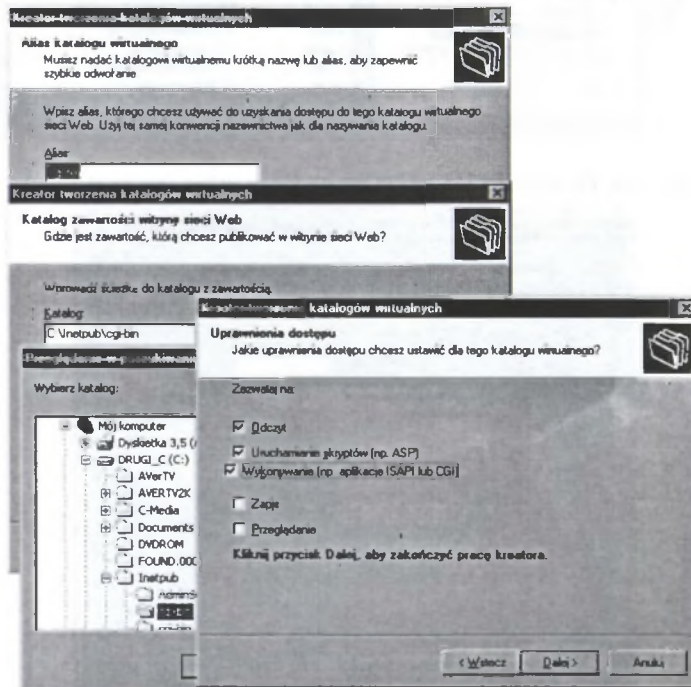
Rys. 100. Konfiguracja sieci w Windows 2000

- Po zainstalowaniu systemu warto sprawdzić (ew. skonfigurować):
- Ogólne ustawieniach serwera – port serwera (rys. 100).
 - Położenie katalogu zawierającego dokumenty serwera (rys. 101).
 - Położenie katalogu zawierającego skrypty cgi (rys. 102).

O ile ustalenie numeru portu⁶ oraz katalogu macierzystego nie powinno nastroczać większych problemów, o tyle ustalenie katalogu, w którym będą znajdowały się skrypt(y) cgi systemu WWW-ISIS, może powodować pewne trudności.



Rys. 101. Konfiguracja IIS (katalog macierzysty)



Rys. 102. Konfiguracja IIS (katalog „cgi”)

⁶ Przeważnie jest to 80, ale w przypadku zainstalowania innego serwera warto zweryfikować tę wartość.

W IIS programy cgi są wykonywane, gdy serwer WWW właściwie interpretuje adres URL zawierający nazwę programu cgi i wszelkie parametry wymagane przez ten program. Jeśli program cgi jest skompilowany w postaci pliku wykonywalnego (.exe), katalogowi, w którym znajduje się ten program, trzeba przyznać uprawnienie wykonywania programów (rys. 101). Atrybut „Wykonywanie” jest konieczny, aby użytkownicy mogli uruchomić zlokalizowane tam skrypty cgi.

Tworzenie katalogu wirtualnego jest proste i polega na wybraniu opcji Create/New, a następnie w menedżerze internetowych usług IIS utworzeniu nowego katalogu wirtualnego, w którym powinno znajdować się odwołanie do utworzonego katalogu dla skryptów cgi z ustalonymi atrybutem Wykonywane (rys. 102).

Serwer Sprite

Serwer Sprite jest oprogramowaniem tzw. open source. Jego częściowej zmiany i przystosowania do potrzeb systemu WWW-ISIS dokonała polska firma CC. W omawianym systemie pełni on rolę serwera obsługującego lokalne bazy danych. Jego przeznaczeniem jest uruchamianie baz danych ISIS znajdujących się lokalnie na płytach CD-ROM. Do poprawnego funkcjonowania systemu konieczne jest jego umieszczenie w osobnym katalogu np.: SPRITE. W katalogu SPRITE umieszczamy zbiory: sprite.exe i killer.exe i programami wsadowymi typu batch oraz z wybranymi komponentami WWW-ISIS.

Samo uruchomienie serwera SPRITE można w pełni zautomatyzować umieszczając w katalogu głównym napędu CD-ROM plik autorun.inf z następującymi poleceniami:

```
[AUTORUN]
OPEN=STARTSRV.BAT
```

Serwer jest automatycznie uruchamiany po włożeniu do napędu przygotowanego krążka CD-ROM. Struktura programu wsadowego STARTSRV.BAT jest dość prosta i polega na:

- uruchomieniu serwera Sprite;
- uruchomieniu przeglądarki WWW (Internet Explorer lub Netscape Navigator) z domyślnym adresem URL (i numerem portu – 9000): <http://127.0.0.1:9000/index.html>



Rys. 103. Uruchomienie serwera Sprite

W momencie wprowadzenia do napędu krążka CD-ROM, przygotowana przez nas baza danych jest automatycznie uruchamiana, a jej adres to <http://127.0.0.1:9000/>. Oczywiście adres 127.0.0.1 jest adresem komputera, na którym uruchamiamy przeglądarkę. Do zatrzymania pracy serwera służy plik wsadowy stopsrv.bat. Nie zaleca się uruchamiania tego pliku, jeśli zamierzamy niebawem powrócić do pracy z bazą.

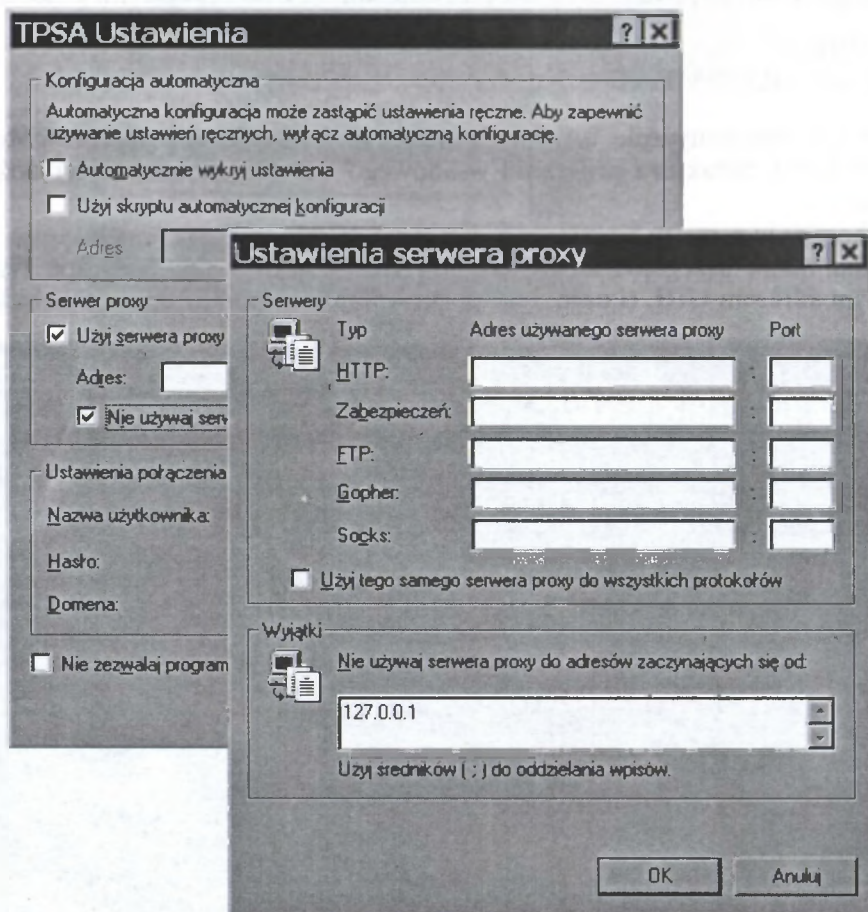
Problemy

Jeśli na danym komputerze nie jest zainstalowana żadna przeglądarka, program rzecz jasna nie uruchomi się i zażąda od nas zainstalowania przeglądarki WWW.

Może się zdarzyć, że procedura automatycznego uruchamiania programu po włożeniu CD-ROM do napędu, nie powiedzie się. Oznaczać to może, że albo nasz napęd jest bardzo wolny, albo opcja automatycznego uruchamiania nie jest włączona w systemie Windows. W takiej sytuacji radzimy uruchomić system ręcznie przez przyciśnięcie ikony z programem wykonywalnym: startsrv.bat.

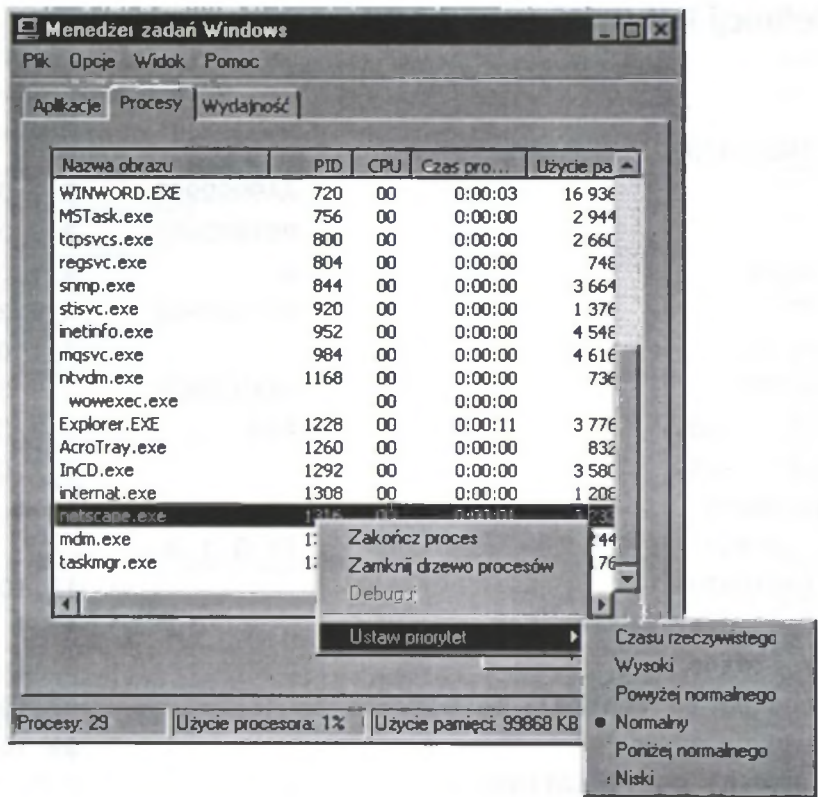
Czasami może okazać się, że system próbuje nawiązać połączenie „Connecting...” jednak bez żadnego rezultatu. Należy sprawdzić wyświetlony w panelu przeglądarki adres, który powinien być: <http://127.0.0.1:9000/>

Jeśli adres jest poprawny, a połączenie nadal nie działa, należy sprawdzić parametry serwera proxy. Powinny być ustawione na bezpośrednie połączenie z Internetem. Jeśli jednak zdecydowaliśmy się na korzystanie z serwera proxy, należy wyłączyć z zasięgu serwera proxy adres 127.0.0.1



Rys. 104. Ustawienia proxy dla serwera Sprite

Port jest ustawiony na 9000, co powinno zapobiec konfliktowi z dowolnym serwerem, działającym na naszym komputerze. Czasami w systemie Windows 2000 może okazać się, że uruchomiona przeglądarka Netscape będzie pracować zbyt wolno, od czasu do czasu przekazując informację o przeciążeniu systemu. Oznacza to, że program Netscape pochłania większość zasobów systemu i tym samym blokuje system. Rozwiązaniem problemu może być przejście do menedżera zadań i zmniejszenie wykorzystania sposobów systemu przez przeglądarkę Netscape.



Rys. 105. Zmniejszenie obciążenia systemu przez przeglądarkę Netscape

Aneks F

Tablice FDT i FST

Tablica Definicji Pól (FDT) bazy AGRIN

SystemRN		1 10 0 0
AGRIŠ RN		90 9 0 0
DATABASE INDICATOR		92 9 0 0
TRN	AA9999999	2 9 3 0
RN	99X999999	3 9 3 0
record status	a	4 1 3 0
affected RN	99X999999	5 9 3 0
transl/generic		6 1 0 0
related RN/TRN	XXX999999	7 9 3 0
primary subj. cat.	A99	8 3 3 0
second. subj. cat.		9 7 0 1
type of document		10 1 1 0
bibliogr. levels	11 3 1 0	
literary indicator		12 10 1 1
literary indicator -Extension		13 1 1 0
Data Entry Status		91 1 0 0
Date of sending to Centre		97 10 0 0
Date of sending with "D"		99 10 0 0
personal name(s) (affiliation)		1100 400 0 0
corporate name(s)		1110 400 0 0
degree		1111 80 0 0
English main title		1200 400 0 0
conference name		1210 200 0 0
conference place		1211 80 0 0
conference date		1213 80 0 0
conference name PFX		1214 80 0 0
conference name SFX		1215 80 0 0
orig. main title		1230 400 0 0
French main title		1232 400 0 0
Spanish main title		1234 400 0 0
edition		1250 80 0 0
report/patent number		1300 80 0 0
secondary number(s)		1310 80 0 0
ISBN/IPC		1320 80 0 0
imprint, place of publication		1401 80 0 0
imprint, publisher		1402 80 0 0
imprint, date of publication		1403 80 0 0

collation	1500	40	0	0
language of text	1600	40	0	1
summary	1601	80	0	1
notes	1610	160	0	0
personal name(s) (affiliation)	2100	400	0	0
corporate name(s)	2110	400	0	0
degree	2111	80	0	0
Engl. main title	2200	400	0	0
conference name	2210	200	0	0
conference place	2211	80	0	0
conference date	2213	80	0	0
conference name PFX	2214	80	0	0
conference name SFX	2215	80	0	0
orig. main title	2230	400	0	0
French main title	2232	400	0	0
Spanish main title	2234	400	0	0
edition	2250	80	0	0
report/patent number	2300	80	0	0
sec. numbers	2310	80	0	0
ISBN/IPC	2320	80	0	0
imprint, place of publication	2401	80	0	0
imprint, publisher	2402	80	0	0
imprint, date of publication	2403	80	0	0
collation	2500	40	0	0
language(s) of text	2600	40	0	1
summary	2601	80	0	1
notes	2610	160	0	0
availability	2611	160	0	0
serial, main title	4230	400	0	0
serial, sec. title	4231	400	0	0
serial, ISSN	4320	80	0	0
serial, date of publication	4403	80	0	0
serial, collation	4500	40	0	0
serial, notes	4610	160	0	0
availability	4611	160	0	0
lang. code (AGROVOC)	5009	2	0	0
lang. code (local index terms)	5029	2	0	0
AGROVOC	5800	600	0	0
AGROVOC after / (not used)	5801	400	0	0
comments/proposed AGROVOC	5810	250	0	0
local index terms	5820	400	0	0
AGROVOC (computer assigned)	5830	700	0	0
AGROVOC (computer codes)	5840	400	0	0
lang. code (1st abstract)	6009	2	0	0
lang. code (2nd abstract)	6019	2	0	0

abstract (1st)	6860 1600 0 0
abstract (2nd)	6861 1600 0 0
reserved	6901 1600 0 0
reserved	6902 1600 0 0
reserved	6903 1600 0 0
institution code	7100 10 0 0
location of document	7200 40 0 1
access number	7300 40 0 1
local RN	7400 9 0 0
title language code	7500 2 1 0
abstract language code	7600 2 1 0
additional abstract	7610 1500 0 0
filled by	7900 20 0 0
date	7910 10 0 0
updated by	7920 20 0 0
date	7930 10 0 0
dummy-0	7997 500 0 0
dummy-1	7998 100 0 1
dummy-2	7999 100 0 1
AUTHORITY TYPE	970 1 0 0
AUTHORITY NAME (J,CA,CO)	971 100 0 0
AUTHORITY NAME (PA)	972 100 0 0
INITIALS	973 10 0 0
AUTHORITY LANGUAGE	977 1 0 0
ISSN in AUTHORITY RECORD	320 30 0 0

Struktura bazy THES

Tablica Definicji Pól (Field Definition Table)

Descriptor	1-st lead. lang.	1 30 0 0
Scope note		2 500 0 1
US	1-st lead. lang.	3 30 0 0
UF	1-st lead. lang.	4 30 0 1
BT		5 30 0 1
NT		6 30 0 1
RT		7 30 0 1
Non-descriptor		8 30 0 0
OTHER	1-st lead. lang.	9 30 0 0
Stopwords	1-st lead. lang.	10 1650 0 1
+US	1-st lead. lang.	11 30 0 1
+UF	1-st lead. lang.	12 30 0 1
Hash flag		19 30 0 1
Top classification name		20 30 0 1
Descriptor 2-nd add. lang.		101 30 0 0
Scope note		102 500 0 0

US	2-nd add. lang.	103	30	0	0
UF	2-nd add. lang.	104	30	0	0
BT		105	30	0	0
NT		106	30	0	0
RT		107	30	0	0
Non-descriptor		108	30	0	0
OTHER	2-nd add. lang.	109	30	0	0
Stopwords	2-nd add. lang.	110	1650	0	1
+US	2-nd add. lang.	111	30	0	1
+UF	2-nd add. lang.	112	30	0	1
... ..					
.. ..					

Tablica selekcji Pól (FST-Field Select Table)

9	0	"T=DSA"D1/ 'T=ALL '
1	0	"E:"V1
8	0	"E:"V8
1	8	'/ES:/',V1
8	8	'/ES:/',V8

Literatura

[BES00]

Hugo Besemer: a Waicent Strategy With Regard To The Development And Support Of Software Applications For Participating Information Centers, 2000-04-07, document w formie elektronicznej.

[3WISIS1]

WWW/ISIS, Technical Reference Manual. Warsaw, 2000. URL <http://www.icie.com.pl/>.

[3WISIS2]

WWW-ISIS: a painless way for publishing ISIS databases on WEB Tutorial v. 1.0. Warsaw, 2001. URL <http://www.icie.com.pl/download>.

[3WISIS3]

WWW-ISIS: a painless way for implementing data entry for ISIS databases on WEB Tutorial, part II v. 1.0. Warsaw, 2001. URL <http://www.icie.com.pl/download>.

[CHE81]

Peter Pin-Shan Chen: a preliminary Framework for Entity-Relationships Models. [W:] (red.) Peter Pin-Shan Chen: Entity-Relationships Approach in Information Modelling and Analysis. ER Institute. 1981.

[DAT00]

C.J. Date: Wprowadzenie do Baz Danych. Warszawa 2000.

[DAT91]

C.J. Date: Entity/Relationship Modeling and the Relational Model [W:] red. C.J. Date: Hugh Darven: Relational Database Writings 1989-1991. 1992.

[HAM78]

Hammer M..M., McLeod D.J.: „The Semantic Data model: a modelling Mechanism for Database Applications” Proc. 1978 ACM SIGMOND International Conference on Management of Data, Austin, Texas (May/June 1978).

[MUR93]

Muraszkiewicz M. Rybiński: Bazy Danych. Warszawa 1993.

[MAN90]

Manual for Preparing Records in Microcomputer-Based Bibliographic Information Systems / Anne Di Lauro Ottawa: IDRC, 1990 (sample CDS/ISIS implementation) Addison, Chris. Installation of CDS/ISIS on a Novell network.

[LAB95]

Di Lauro, Anne; Brandon, Ed MIBIS guide: manual for preparing record in microcomputer-based bibliographic information systems. IDRC, Ontario, Canada. 1995, 2nd ed. 355 p. ISBN 0-88936-559-8.

[HBA94]

Hopkinson, Alan; Buxton, Andrew: The CDS/ISIS handbook. Library Association Publishing. . London 1994.

[HOP91]

Hopkinson, Alan. "CDS/ISIS for microcomputers as an interim library automation system." Online Information 91 : 15th International Online Information Meeting, Proceedings, London 10-12 December 1991/ edited by David I. Raitt, 499-504. Oxford and New Jersey: Learned Information (Europe), 1991.

[HOP94]

Hopkinson, Alan. International Information System on Cultural Development : CDS/ISIS Model Data Base : Manual and Accompanying Diskette, iii, 190 p. + 1 floppy diskette 5.25" PGI-93/WS/, 16. Paris: General Information Programme and UNISIST, UNESCO, 1994.

[HOP94a]

Hopkinson, A. CDS/ISIS as a tool for implementing UNIMARC. In: UNIMARC and CDS/ISIS : proceedings of the workshops held In Budapest and Barcelona. 1993, Munich, Saur, 1994, 37-46.

[SPO99]

Sportack M: Sieci komputerowe, księga eksperta. Warszawa 1999.

[SHE95]

Sheldon T.: Wielka encyklopedia sieci komputerowych. Wrocław 1995.

[DNO96]

Dobosz J., Nowicki Z. M., Ostrowski M., Romański S., Sadoch A., Struk W., Sulej M.: Micro CDS/ISIS, Pytania i odpowiedzi. Warszawa 1996.

[FAN991]

Faber R., Nowicki: Użytkowanie pakietu Mikro CDS/ISIS. Warszawa 1991.

[BOH01]

Buxton A., Hopkinson A.: The CDS/ISIS for Windows Handbook.. Paris 2001.

[NOW92]

Nowicki, Z. M. How to utilize advanced features of micro CDS/ISIS, Preliminary unedited version ed. ITC/DPMD/EDP, 6, Rev. 2. Geneva: International Trade Centre, 1992.

[ISIS91]

Mini Mikro CDS/ISI (Wersja 2.1) Dokumentacja Użytkownika. Warszawa 1991.

Spis rysunków, tabel i tablic

Rysunki

Rys. 1. Ogólny plan podręcznika	10
Rys. 2. Diagram E-R: „Pracownik pracuje nad projektem”	15
Rys. 3. Diagram E-R: „Student zalicza wykład”	15
Rys. 4. Diagram E-R: Przykład zależności egzystencjalnej WYDZIAŁY UCZELNI ...	17
Rys. 5. Diagram E-R: Zależność identyfikacyjna WYKŁADOWCA_WYKŁAD	18
Rys. 6. Model E-R: „Student zalicza wykład” (skrótowy)	19
Rys. 7. Diagram E-R: „Student zalicza wykład” (pełny)	19
Rys. 8. Diagram E-R: dla relacji „Student zalicza wykład”	20
Rys. 9. Implementacja modelu E-R: „Student zalicza wykład” (z przykładowymi danymi) w typowo relacyjnym systemie zarządzania bazą danych Access	21
Rys. 10. Diagram E-R: „...artykuł opublikowany w...”	22
Rys. 11. Diagram E-R: relacja trójargumentowa „WYPOŻYCZANIE KSIĄZEK” ..	22
Rys. 12. Diagram E-R: relacja binarna ze słabą encją	23
Rys. 13. Diagram E-R: relacja binarna ze słabą encją (z atrybutami i kluczami)	23
Rys. 14. Diagram E-R: „Sprzedający kupujący”	25
Rys. 15. Relacje między zbiorem głównym a skorowidzem	33
Rys. 16. Relacje między zbiorem głównym a zbiorem odwróconym	33
Rys. 17. Diagram E-R: „...artykuł opublikowany w...”	36
Rys. 18. WWW-ISIS – ogólny przegląd	41
Rys. 19. Działanie systemu	42
Rys. 20. Wstępny plan implementacji systemu – sieć Internet/Intranet	44
Rys. 21. Ustalenie parametrów katalogu macierzystego	46
Rys. 22. Ustalenie katalogu skryptów cgi	46
Rys. 23. Problemy z WWW-ISIS – brak dostępu do serwera	48
Rys. 24. Problemy z WWW-ISIS – brak zbiorów HTML	48
Rys. 25. Problemy z WWW-ISIS – brak prawej ramki	49
Rys. 26. Logo BIREME	50
Rys. 27. Usuwanie logo BIREME w Windows 9x	51
Rys. 28. Usuwanie logo BIREME w Windows 2000	51
Rys. 29. Powitalny ekran aplikacji DEMO systemu WWW-ISIS	52
Rys. 30. Aplikacja DEMO – katalogi	53
Rys. 31. Opcja przeglądania danych w tezaurusie	55
Rys. 32. Formularz wyszukiwawczy: operatory AND, OR łączą terminy wewnątrz pola, zaś operatory AND, OR, NOT łączą pola.	56
Rys. 33. Korzystanie ze słownika (przycisk – List)	56
Rys. 34. Reedycja zapytania w języku wyszukiwawczym ISIS (EDIT QUERY)	57
Rys. 35. Historia wyszukiwań (Opcja „History”)	57
Rys. 36. Wyszukiwanie za pomocą tezaury	58
Rys. 37. Schemat interakcji w interfejsie wprowadzania danych	59
Rys. 38. Wprowadzanie nowego rekordu	60
Rys. 39. Uaktualnianie rekordu	61
Rys. 40. Kopiowanie rekordu	62
Rys. 41. Ustawienia „cookies” w Netscape Navigator	63
Rys. 42. Strona rejestracyjna systemu wprowadzania danych WWW-ISIS	63
Rys. 43. Ustawienia „cookies” w Internet Explorer	64
Rys. 44. Przykładowe pole interfejsu do wprowadzania danych WWW-ISIS	65
Rys. 45. Tablica podglądowa dla pola „Corporate author”	66
Rys. 46. Grupy powtarzalne w interfejsie wprowadzania danych WWW-ISIS	66
Rys. 47. Omówienie interfejsu wyszukiwawczego WWW-ISIS	69
Rys. 48. Schemat implementacji systemu WWW-ISIS (wyszukiwanie danych)	71

Rys. 49. Pole typu kombi w interfejsie wyszukiwawczym	93
Rys. 50. Zastosowanie formatów wyświetlania i sortowania	94
Rys. 51. Przykład pola kombi do wyboru formatów wyświetlania danych	95
Rys. 52. Nawigacja w obrębie tezaury	114
Rys. 53. Formularz z polem ustalającym zakres wyszukiwań	118
Rys. 54. Wyszukiwanie pełnotekstowe	118
Rys. 55. Grupy And-Or	119
Rys. 56. Omówienie systemu wprowadzania danych WWW-ISIS	123
Rys. 57. Schemat implementacji systemu WWW-ISIS (wprowadzanie danych)	124
Rys. 58. Pole tekstowe	132
Rys. 59. Wielowierszowe pole tekstowe	132
Rys. 60. Pole typu kombi pojedynczego wyboru	132
Rys. 61. Pole kombi wielokrotnego wyboru	133
Rys. 62. Pole typu zaznacz (check)	133
Rys. 63. Pole typu przełącznik (radio)	133
Rys. 64. Struktura zbioru entry.def (tablica definicji arkuszy wprowadzania danych)	134
Rys. 65. Jedno i wielowierszowe pola tekstowe (typ: 1, 2)	139
Rys. 66. Przykład listy odsyłaczy (pól wyboru)	141
Rys. 67. Przykład wielowierszowego pola tekstowego	143
Rys. 68. Pole „Corporate author” z tablicą odnośników	143
Rys. 69. Wprowadzanie nowego rekordu do tabeli odnośników	145
Rys. 70. Lista z polami typu „zaznacz” w tezaurysie	147
Rys. 71. Pola typu kombi pojedynczego i wielokrotnego wyboru (typ 3 i 4)	148
Rys. 72. Pola typu „zaznacz” i „przełącznik” (typ 5 i 6)	149
Rys. 73. Komunikat o błędzie (nie można zapisać rekordu)	153
Rys. 74. Ostrzeżenie (można zapisać rekord do bazy)	154
Rys. 75. Strona wywoławcza pozwalająca na rejestrację w systemie (login)	160
Rys. 76. Strona wyboru rekordów do uaktualnienia	163
Rys. 77. Ustawienia pamięci dla programu ISIS (wersja dla DOS)	184
Rys. 78. Przykładowe menu CDS/ISIS dla DOS	190
Rys. 79. Przykładowe menu CDS/ISIS dla Windows	191
Rys. 80. Obsługa zbioru głównego	203
Rys. 81. Ustalenie parametrów eksportu	204
Rys. 82. Ustalenie parametrów importu	204
Rys. 83. Program FANGORN – ustalenie języka konwersacji	205
Rys. 84. FANGORN – tworzenie zbioru konfiguracyjnego	206
Rys. 85. Podanie nazwy zbioru konfiguracyjnego	209
Rys. 86. Przeprowadzenie konwersji	209
Rys. 87. Importowanie danych	210
Rys. 88. Architektura klient-serwer	212
Rys. 89. Architektura klient – serwer, system WWW	213
Rys. 90. Zasada działania protokołu IP	213
Rys. 91. Przyporządkowanie nazw DNS i adresów IP	214
Rys. 92. Instalacja protokołu TCP/IP	214
Rys. 93. Konfiguracja maski podsieci i adresu IP	215
Rys. 94. Ustawienia bramy w protokole TCP/IP	215
Rys. 95. Określenie adresów IP serwerów DNS	216
Rys. 96. Powitalna strona administratora	217
Rys. 97. XITAMI panel administracyjny – główne ustawienia serwera WWW	217
Rys. 98. Instalacja IIS (Internetowe Usługi Informacyjne)	218
Rys. 99. Dostęp do IIS przez menedżera usług internetowych	219
Rys. 100. Konfiguracja sieci w Windows 2000	219
Rys. 101. Konfiguracja IIS (katalog macierzysty)	220
Rys. 102. Konfiguracja IIS (katalog „cgi”)	220
Rys. 103. Uruchomienie serwera Sprite	221
Rys. 104. Ustawienia Proxy dla serwera Sprite	222
Rys. 105. Zmniejszenie obciążenia systemu przez przeglądarkę Netscape	223

Tabele

Tabela 1. Przykłady określenia podstawowych pojęć E-R	16
Tabela 2. Struktura danych dla modelu E-R: „STUDENT ZALICZA WYKŁAD” ..	20
Tabela 3. Układ danych dla przykładowego modelu „CZYTELNIK WYPOŻYCZA KSIĄŻKĘ”	24
Tabela 4. Tabela porównawcza dla bazy DEMO	77
Tabela 5. Tablica wyboru pól (FST) dla bazy AGRIN	83
Tabela 6. Tworzenie indeksów dla wyszukiwania uproszczonego (baza AGRIN) ...	87
Tabela 7. Tworzenie indeksów dla wyszukiwania zaawansowanego (baza AGRIN) ..	88
Tabela 8. FST (WWW-ISIS) dla bazy AGRIN	88
Tabela 9. Omówienie tablicy wyszukiwań WWW-ISIS dla systemu DEMO	90
Tabela 10. Przykładowy format dla bazy DEMO (pole/zawartość)	103
Tabela 11. Tablica przeglądania (browse.def.eng) WWW-ISIS dla bazy THES	112
Tabela 12. Przykład zbioru dbn.def	113
Tabela 13. Krok 2 w projektowaniu systemu wprowadzania danych WWW-ISIS ...	127
Tabela 14. Parametry dla pól (arkusz wprowadzania danych)	139
Tabela 15. Operacje wykonywane przez funkcje ON_NEW i ON_UPDATE	156
Tabela 16. Fragment formatu zatwierdzania danych.	168
Tabela 17. Zbiory HTP w WWW-ISIS.	174
Tabela 18. Parametry przykładowego zbioru konfiguracyjnego isis3w.exe.cfg.	178
Tabela 19. Omówienie wszystkich parametrów zbioru konfiguracyjnego (isis3w.exe.cfg)	180
Tabela 20. Fragment tablicy definicji pól w CDS/ISIS.	187
Tabela 21. Przykładowy zbiór konfiguracyjny lang.par dla programu FANGORN ..	206
Tabela 22. Przykładowe dane w arkuszu Excella	207

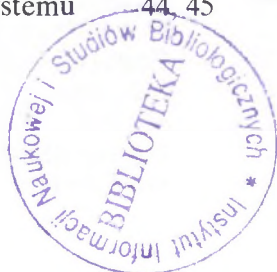
Tablice

Tablica 1. Ważniejsze skróty i oznaczenia przyjęte w książce	11
Tablica 2. WWW-ISIS na tle innych systemów obsługujących bazy ISIS	28
Tablica 3. Format ESMON.PFT dla bazy DEMO	99
Tablica 4. Parametry arkusza wprowadzania danych (isis_worksheet)	135
Tablica 5. Typy elementów wprowadzania danych	138
Tablica 6. Parametry dla tablic odnośników (interfejs wprowadzania danych)	141
Tablica 7. Fragment formatu odpowiedzialnego za przeglądanie rekordów.	164

Skorowidz ważniejszych terminów

adres IP użytkownika bazy	155,157	isis_worksheet	135-136
adres URL		Lookup_choice	142
de_worksheet	135	lookup_type	141
page_footer	106	Path	141
page_header	106	Pfx	141
record_header	106	Startterm	141
search_type=add_thes_word	104	Table	142
search_type=browse	104	error.http	174
search_type=ef_copy	163	FANGORN	205-210
search_type=ef_login	169	konwersja danych	208
search_type=ef_logout	169	tworzenie tabeli	207
search_type=ef_modify	164	ustalenie parametrów	205-206
search_type=link	103	wejściowy zbiór tekstowy	208
search_type=query	103	ISIS	184-204
architektura klient-serwer	212	eksport	203
browse.def	74, 75, 109	grupy powtarzalne	195
binding	111	import	203
format	111	ISO2709	203
prefix	110-111	język formatowania	192-198
typ rekordu	110	zbiór ANY	189
browse.http	174	zbiór główny	188
CDS/ISIS	27-39, 184-204	zbiory definiujące bazę	187
syspar.par	185-186	ISIS formaty	<i>patrz</i> ISIS język formato-
database.def	90-98	wiania	
definicje typu rekordu	91	ISIS FST	188-189
definicje wyboru pól	92	techniki selekcji	188-189
format	95-96	ISIS język formatowania	192-198
restricted_terms	96	funkcja VAL	197
sort	94	literały	194-196
DNS	214	instrukcja if	197
hosts	216	ISISREF funkcja	115
e_browse.http	174	isis3w.exe.cfg	178-183
e_forms.http	174	AndOr	180
encja	16	BrowseDefinitionFile	180
entry.def (parametry)	132-149	BrowsePaging	180
Create	142	CookieDomain	180
Create_type	142	CookiePath	180
Dbn	141	DBNFile	181
Dlmtr_l	142	DB_name	180
Dlmtr_r	142	DebugLevel	180
Fmt	141	DebugLogFile	180
Gen_occ	142	DefaultFormat	180
Init_pft	142	EntryDefinitionFile	180
isis_field	136-137	FieldAndOr	180
isis_page	136	Format	180
		HighlightSearchTerms	180
		HttpFilePath	180

IncludeFilePath	181	system wprowadzania danych	123-175
IsisDataDir	181	etapy implementacji	124
IsisDocLocation	181	funkcja ADD	156
IsisDocLocationOcc	181	funkcja DEL	156
IsisForceType	181	funkcja ON NEW	155-157
IsisMultiDb	181	funkcja ON UPDATE	155-157
IsisMultiDbFile	181	funkcja REPL	156
IsisMultiDbUser	181	funkcja SR	156
IsisTempFileLife	181	on_new_pft	135
IsisWorkDir	181	plan działania	125-126
LatinConvOn	181	pola tekstowe	132, 138, 139
LatinConvTerms	181	pole kombi pojedynczego	
ListCheckBox	181	wyboru	132, 138, 148
PagingMaxPages	181	pole kombi wielokrotnego	
PagingMaxRows	181	wyboru	133, 138, 148
QueryMaxRows	182	pole przełącznik	133, 138, 149
RestrictedTermMinLength	182	pole zaznaczyć	133, 138, 149
SearchDefinitionFile	182	użytkownik PUBLIC	159
SearchPaging	182	tablica definicji przeglądania danych	
Sort	182	<i>patrz</i> browse.def	
TermSeparator	182	tablica definicji wyszukiwań <i>patrz</i> data-	
podtyp	16	base.def	
protokół TCP/IP	213-216	tezaurus	58
adres IP	214	indeks alfabetyczny	111, 114
konfiguracja	215	indeks KWOC	111, 114
przeglądarka WWW	45	thes.http	174
REF funkcja	115, 196	uaktualnianie <i>patrz</i> wprowadzanie	
relacja	16	danych	
result.http	174	view.http	174
search.http	174	wirtualne pola	
Serwer IIS	218-221	9950	155, 166
konfiguracja	219	9960	155, 166
ustawienia	220	9961	155, 166
problemy	220-221	9970	155, 166
Serwer Sprite	221-223	9991	105, 115
instalacja	221	9999	105, 115
konfiguracja	222	własność	16
problemy	222-223	wyszukiwanie danych	
Serwer XITAMI	216-218	database.def	90
instalacja	216	eksport	116
konfiguracja	216-218	metoda skrócona	71, 75-84
problemy	218	metoda standardowa	84-107
Serwery WWW	43, 216	zapamiętywanie zapytań	116
plan implementacji systemu	44, 45	związek	<i>patrz</i> „relacja”



22502 + CD
M.2

**Seria wydawana z inicjatywy
INSTYTUTU INFORMACJI NAUKOWEJ
I STUDIÓW BIBLIOLOGICZNYCH
UNIwersYTETU WARSZAWSKIEGO
oraz
WYDAWNICTWA SBP**