

Bohdana Stoklasová, Miroslav Bareš, Anna Nerglová



**TECHNOLOGIE RETROKON
JAKO KOMPLEXNÍ NÁSTROJ PRO DIGITALIZACI
A ZPŘÍSTUPNĚNÍ KATALOGŮ, KARTOTÉK
A SOUPISŮ PROSTŘEDNICTVÍM INTERNETU**

Bohdana Stoklasová, Miroslav Bareš, Anna Nerglová

**TECHNOLOGIE RETROKON
JAKO KOMPLEXNÍ NÁSTROJ PRO DIGITALIZACI
A ZPŘÍSTUPNĚNÍ KATALOGŮ, KARTOTÉK
A SOUPISŮ PROSTŘEDNICTVÍM INTERNETU**



Praha 2004
Národní knihovna České republiky

Katalogizace v knize – Národní knihovna České republiky

Stoklasová, Bohdana, 1953-

Technologie RETROKON jako komplexní nástroj pro digitalizaci a zpřístupnění katalogů, kartoték a soupisů prostřednictvím internetu / Bohdana Stoklasová, Miroslav Bareš, Anna Nerglová. – Praha : Národní knihovna České republiky, 2004. – 34 s.

ISBN 80-7050-458-7

004.057.6-026.452 * 002.2:004 * 025.32/.33 * 004 * 021.6 * (437.3)

- RETROKON
- retrospektivní konverze – Česko
- digitalizace dokumentů
- katalogizační záznamy
- informační technologie – Česko
- kooperace knihoven – Česko
- informační publikace

004 - Počítačová věda. Výpočetní technika. Informační technologie

I. Název

II. Bareš, Miroslav, 1963-

III. Nerglová, Anna, 1979-

© Národní knihovna České republiky, 2004

© Bohdana Stoklasová

© Miroslav Bareš

© Anna Nerglová

© Martin Zhouf

ISBN 80-7050-458-7

Obsah:

1 Úvod	5
2 Jádru technologie RETROKON	7
2.1 Skenování a zpřístupnění	7
2.2 Přepis	15
2.3 Strukturování (tagování)	15
2.4 Revize	17
3 RIS – systém pro řízení retrospektivní konverze	18
3.1 Prohlížení záznamů.....	18
3.2 Správa objednávek.....	21
3.3 Označování záznamů	23
3.4 Uložky	25
3.5 Řízení retrospektivní konverze	26
4 Závěr	33
5 Další informace a kontakty	34

RETROKON

101010010101
10101011010101
101001001010101
11011010001011
101101001001
001011011011010
10101101101101
110101001011
101010100101
101010100101
10101011010101
10101011010101
10100100010111
11011010001011
1011010110011
0010110111010
1010110110101
01101011011010
1010100101011
10101011010111
1000000000011
1011010110011
0010110111010
1010110110101
01101011011010
1010100101011
10101011010111
10100100010111

101010001011
1010110011
.....

10010100101011
10101011010111
10100100010111

1 Úvod

Základ technologie RETROKON vznikl před deseti lety v souvislosti se zahájením retrospektivní konverze (převodu lístkových katalogů a bibliografických soupisů z původní papírové do digitální podoby) v Národní knihovně České republiky. Jedná se o originální českou technologii, která je výsledkem spolupráce Národní knihovny ČR (knihovnická část) a firmy Comdat (vývojová a programátorská část a zajištění provozu). Od počátku se jednalo o komplexní technologii, jejíž modulárnost umožňovala postup po etapách. Za uplynulých deset let bylo pomocí této technologie zpřístupněno v České republice zhruba deset miliónů bibliografických a katalogizačních záznamů různých knihoven, což znamenalo zásadní zlom ve smyslu otevření a širokého zpřístupnění jejich fondů nejen místním uživatelům, ale prostřednictvím internetu i uživatelům vzdáleným.

Desetiletá praktická zkušenost a její uplatnění v různých knihovnách ČR umožnily technologii dokonale prověřit, odladit a optimalizovat. V průběhu deseti let se dosti zásadně změnil pohled na možnosti a meze využití jednotlivých etap technologie RETROKON. Na jedné straně došlo k dramatickému nárůstu přenosových rychlostí v síti internet, které umožnily široké využití obrazových katalogů. Na druhé straně se nenaplnily očekávání a plány týkající se financování úplné retrospektivní konverze katalogů českých knihoven (ve smyslu jejich úplného převodu do strukturované podoby umožňující vícehlediskové prohledávání v online katalogu), čímž došlo k podstatnému posunu časových horizontů. Byl-li při vývoji technologie před deseti lety akcentován především modul strukturování (tagování), v průběhu uplynulých deseti let se ukázal jako stejně důležitý modul zpřístupnění naskenovaných obrazových katalogů. Tento modul byl podstatně zdokonalen a obohacen o možnost objednávání dokumentů. Z modulu KATIF, který sloužil ke zpřístupnění naskenovaných katalogů Národní knihovny ČR, vznikl postupně modul NRIS (Národní informační systém pro retrospektivní konverzi), v jehož rámci jsou zpřístupněny katalogy řady českých knihoven (souborný katalog českých knihoven v obrazové podobě) i informace o postupu retrospektivní konverze v českých knihovnách. Omezená dostupnost finančních prostředků si vynutila zejména ve velkých knihovnách postup po etapách a vyřazení záznamů méně využívaných dokumentů z první fáze úplné retrospektivní konverze. Z požadavku na snadnou a spolehlivou evidenci postupu retrospektivní konverze a na management velkých týmů pracovníků v dlouhodobém časovém horizontu v rámci jedné instituce nebo skupiny institucí se zrodil modul Informační systém pro retrospektivní konverzi (RIS), který byl na základě požadavku na přehled o retrospektivní konverzi a její management v celonárodním měřítku částečně transformován v Národní informační systém pro retrospektivní konverzi.

Cílem této publikace je představit technologii RETROKON v celé její komplexnosti pracovníkům knihoven, kteří většinou znají pouze některé z modulů, dále pak pracovníkům archivů, muzeí, galerií a dalších paměťových institucí.

2 Jádru technologie RETROKON

Jádru technologie RETROKON se skládá ze čtyř po sobě následujících modulů (fází), které jsou stejně dobře využitelné jako celek i jednotlivě. Použitelnost jádra technologie RETROKON je univerzální – bez ohledu na charakter a velikost instituce. V následujícím textu budou popsány jednotlivé moduly (fáze) v celkovém kontextu i separátně.

1. Skenování a zpřístupnění
2. Přepis
3. Strukturování (tagování)
4. Revize

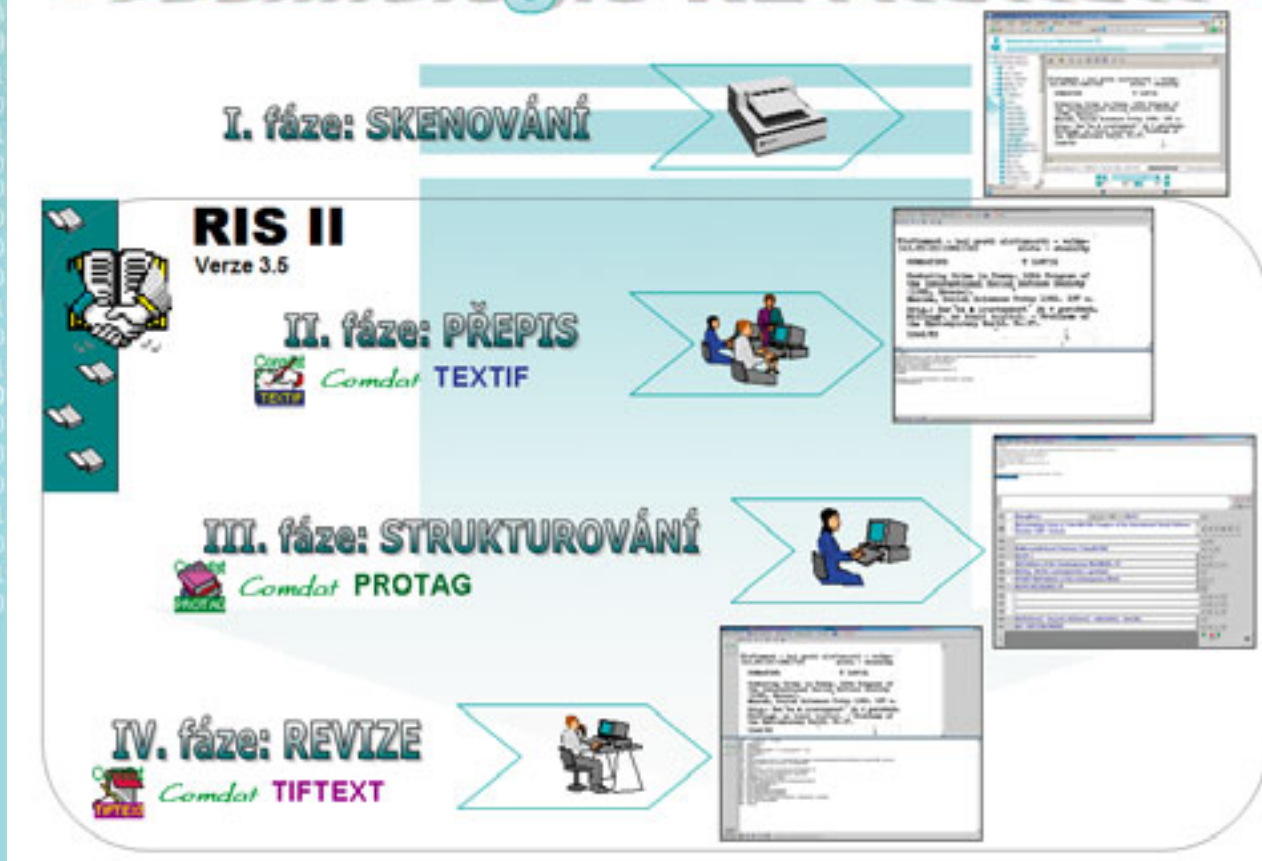
Jednotlivé fáze včetně příslušných programů jsou vyobrazeny na obrázku č. 1. V první etapě je pro skenování využíván program Comdat Artif. V následujících etapách jsou využívány postupně programy Comdat Textif, ProTag a TifText. Tyto čtyři programy tvoří jádro technologie RETROKON. Jádro společně s programy RIS a NRIS (podrobnější popis následuje v dalších kapitolách) tvoří celý komplex technologie RETROKON.

2.1 Skenování a zpřístupnění

Skenování

První fází jádra technologie RETROKON je pořízení obrazové kopie katalogu skenováním a jeho následné zpřístupnění na internetu. Skenování katalogů na vysokokapacitních skenerech umožní pořídit obrazovou kopii katalogu v relativně krátkém časovém období týdnů až měsíců (například skenování generálních katalogů Národní knihovny ČR probíhalo v několika směnách tempem 20 000 lístků denně). Doba skenování závisí na rozsahu katalogů a jejich charakteru (u vysoce využívaných katalogů je žádoucí minimalizovat dobu a rozsah jejich odstavení) a množství dostupných finančních prostředků. Skenování se provádí v místě instalace katalogu nebo je katalog postupně převážen z knihovny na místo skenování a následně vrácen zpět v původním stavu (obvyklý způsob zpracování).

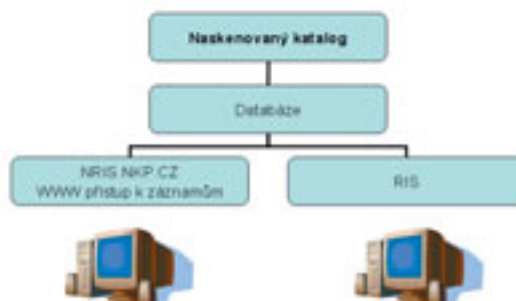
Technologie RETROKON



Obr. 1

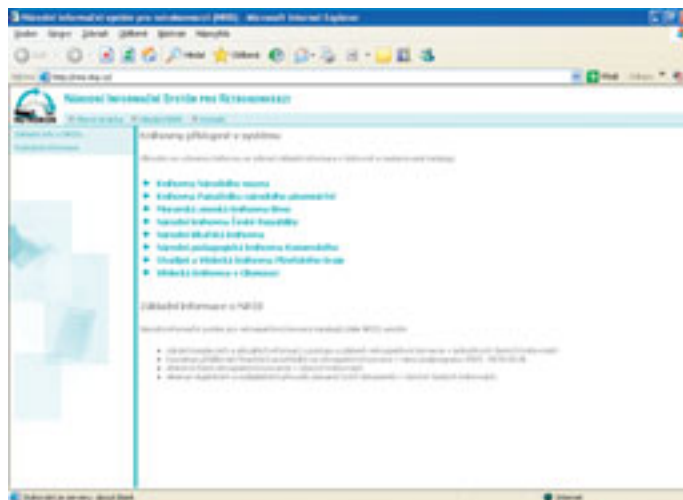
Zpřístupnění

Naskenované lístky vytvoří základ pro zpřístupnění na internetu a zároveň pro informační systém RIS. Informace o záznamech a jejich uložení jsou ukládány do databáze, která následně řídí přístup k záznamům (viz obr. 2).



Obr. 2

Pro snadný a přehledný přístup k záznamům prostřednictvím internetu byl vytvořen systém NRIS, který umožňuje prohlížení libovolného katalogu zařazeného v systému a zprostředkovává informace o průběhu retrospektivní konverze záznamů v daném katalogu. Systém je dostupný na internetové adrese <http://nris.nkp.cz> (viz obr. 3).

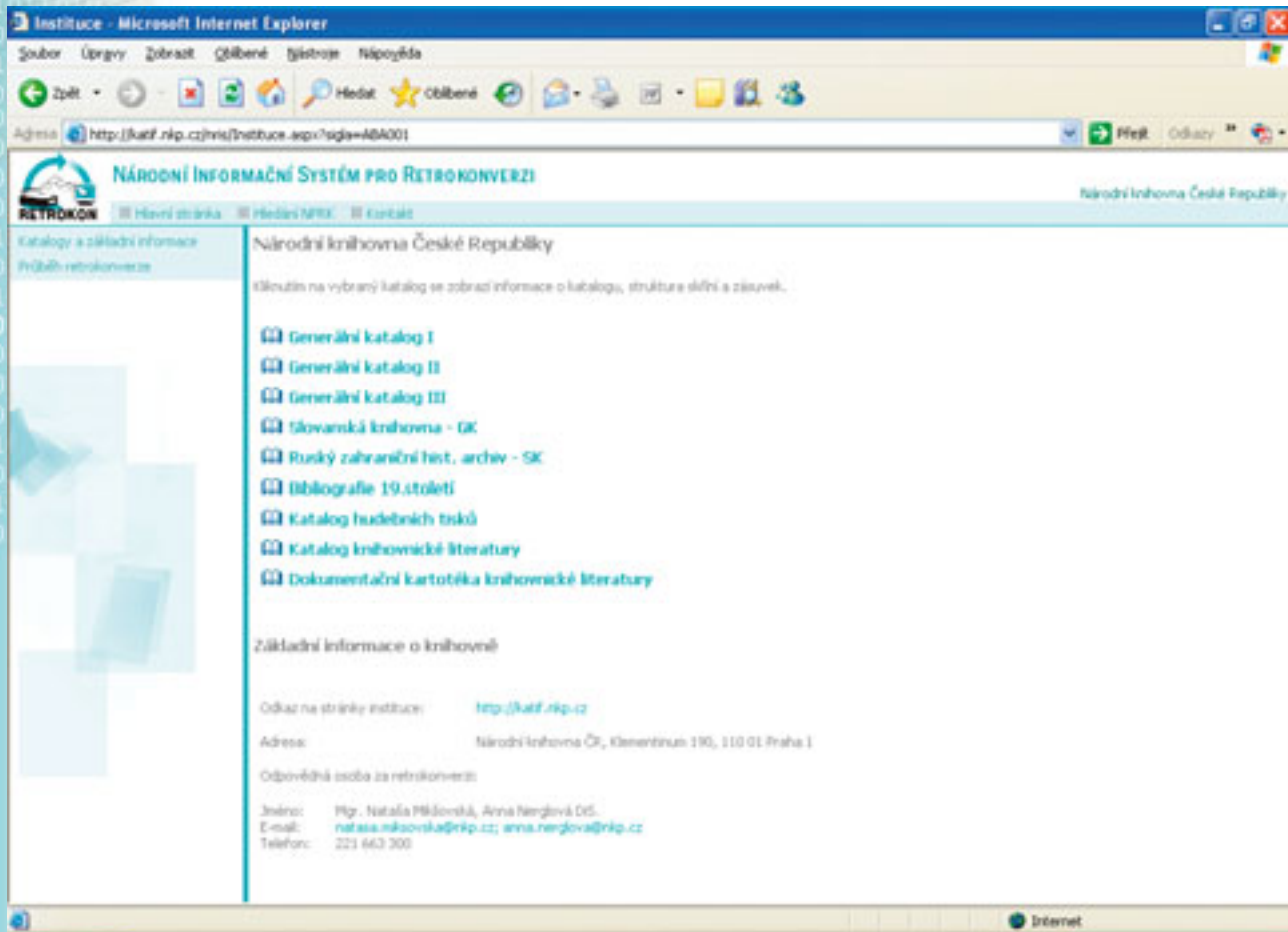


Obr. 3

Úvodní stránka nabízí seznam zúčastněných knihoven a zároveň souhrn základních informací o systému NRIS a jeho významu. Počet knihoven zařazených do systému není limitován a je průběžně rozšiřován. Součástí systému se může stát libovolná knihovna ČR, která o účast projeví zájem. Na konci roku 2004, kdy je připravována tato publikace, je v systému zařazeno 8 knihoven se svými 44 katalogy a 9 364 366 záznamy.

Po výběru kterékoli z nabízených knihoven získává uživatel přehled katalogů zveřejněných danou knihovnou (viz obr. 4).

Obr. 4



System umožňuje prezentovat informace o průběhu retrospektivní konverze nejen v dané knihovně, ale i v příslušném katalogu nebo skupině záznamů (viz obr. 5). Tato služba využívá propojení mezi systémem RIS pro řízení průběhu retrospektivní konverze a systémem NRIS pro zpřístupnění naskenovaných katalogů.

Obr. 5

NÁRODNÍ INFORMAČNÍ SYSTÉM PRO RETROKONVERZI
Národní knihovna České republiky

Katalogy a základní informace
Průběh retrokonverze

Národní knihovna České republiky

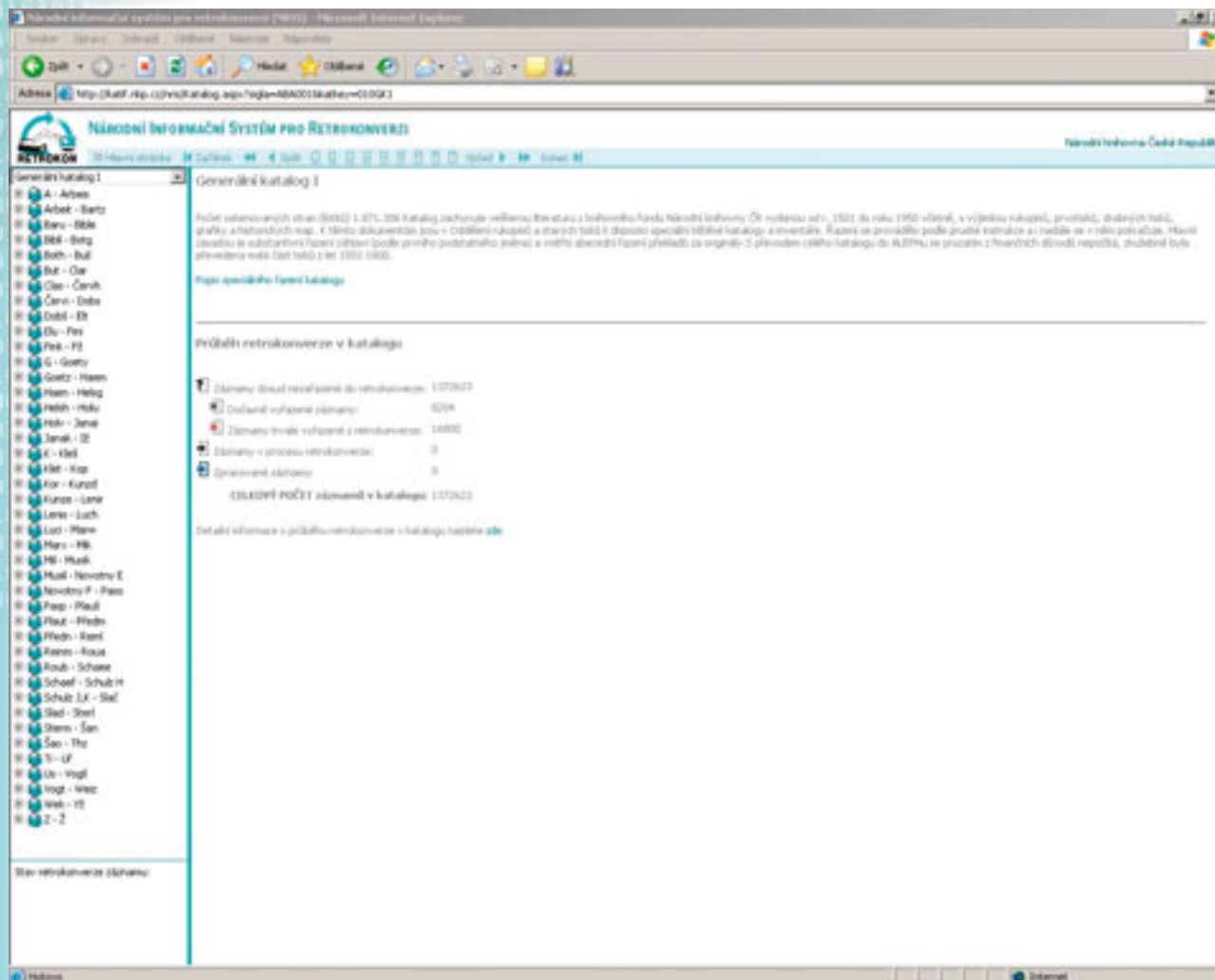
Užitím na vybraný katalog se zobrazí informace o průběhu retrokonverze v různých záznamech.

	↑	↓	↔	↔	↔	↔	Číslem
Generální katalog I	1372797	8204	16793	0	0	0	1372797
Generální katalog II	1240380	393341	270956	213151	5099	0	1458632
Generální katalog III	66487	0	0	0	0	0	66487
Slovenská knihovna - GK	419750	21090	58519	149576	552	0	569878
Ruský zahraniční řet. archív - SK	38629	0	0	0	0	0	38629
Bibliografie 19. století	64222	0	0	0	0	0	64222
Katalog hudebních tisků	90335	0	0	0	0	0	90335
Katalog knihovnické literatury	11910	0	0	0	0	0	11910
Dokumentační kartotéka knihovnické literatury	59572	0	0	0	0	0	59572

(údaje o výsledcích projektu:

Základem zpřístupnění katalogů je obrazová prezentace jejich záznamů. Po výběru katalogu se uživateli zobrazí stromová struktura řazení záznamů v katalogu a základní informace o vybraném katalogu včetně souhrnných informací o stavu retrospektivní konverze – v případě, že je příslušná knihovna uvádí (viz obr. 6).

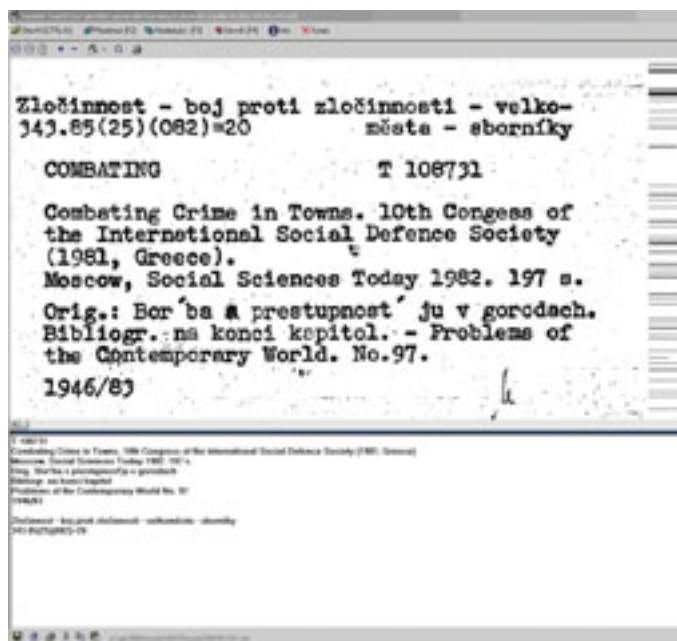
Obr. 6



Sekvenční procházení katalogu umožňuje prohlížení katalogu po jednotlivých záznamech (listování) nebo rychlý posun po desítkách záznamů s možností volby přibližného umístění záznamu. Na pravé straně obrazovky jsou znázorněna pořadová čísla záznamů z vybrané oblasti. Kliknutím na konkrétní číslo se uživateli zobrazí odpovídající záznam. Uživatel může také měnit kvalitu zobrazení záznamů ve dvou úrovních. Nižší kvalita zobrazení umožňuje rychlejší procházení katalogu a vyšší kvalita výrazně zlepšuje čitelnost záznamů. Uživatelské rozhraní slučuje výhody všech obdobných systémů. Výhody spočívají ve sloučení přehledné a jednoduché orientace v katalogu s rychlým a jednoduchým přístupem k záznamům. Důležitý je také snadný přechod mezi katalogy a knihovnami, doplněný o možnost řízení kvality zobrazení a možnost objednávání.

Systém umožňuje objednávat díla z naskenovaných záznamů. Na pozadí vybraného záznamu vyplní čtenář objednávací formulář (viz obr. 8), který je odeslán ke zpracování do knihovny. Objednávky jsou vyřizovány systémem RIS (podrobný popis je uveden v kapitole 3.2 RIS – Správa objednávek). O stavu zpracování objednávky je čtenář průběžně informován dvěma způsoby, prostřednictvím e-mailu nebo www rozhraní.

Obr. 8



Obr. 9

Klientem systému může být libovolný internetový prohlížeč. Řešení není určeno pouze pro klienty pracující pod operačním systémem Microsoft Windows.

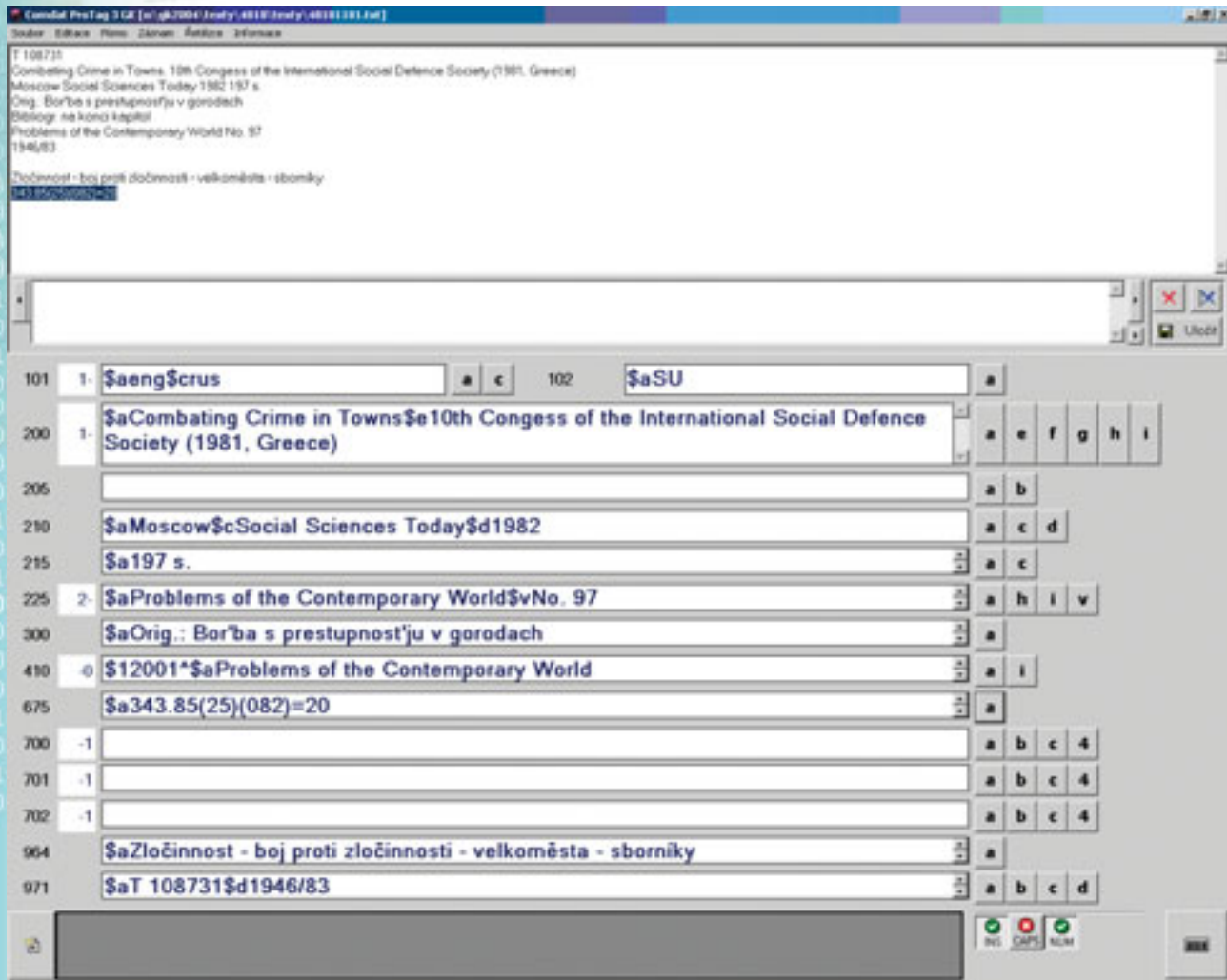
Naskenování katalogů a jejich zpřístupnění v obrazové podobě je rychlé a v porovnání s úplnou retrospektivní konverzí katalogů i levné. Může sloužit jako definitivní varianta i jako varianta přechodná v případech, kde je plánována (a s ohledem na vysokou míru využití i dostatečné finanční zabezpečení i reálná) úplná retrospektivní konverze katalogů. Výsledkem úplné retrospektivní konverze je text strukturovaný v souladu s mezinárodními standardy, jejichž aplikace umožňuje využití v různých systémech a zemích a vícehlediskové prohledávání podle různých selekčních kritérií. Naskenovaný katalog velmi usnadňuje a racionalizuje úplnou retrospektivní konverzi, ať již je prováděna přímo v rámci integrovaného knihovnického systému nebo s využitím dalších modulů technologie RETROKON, které jsou popsány dále.

2.2 Přepis

Cílem této etapy je převedení obrazové podoby záznamů do nestrukturovaného textového tvaru. Přepis může být prováděn dvěma způsoby. Automatickým OCR (Optical Character Recognition) přepisem je možné zpracovat kvalitní strojem psané záznamy. Pro ruční přepis méně kvalitních předloh byl vyvinut program Comdat Textif (viz obr. 9).

2.3 Strukturování (tagování)

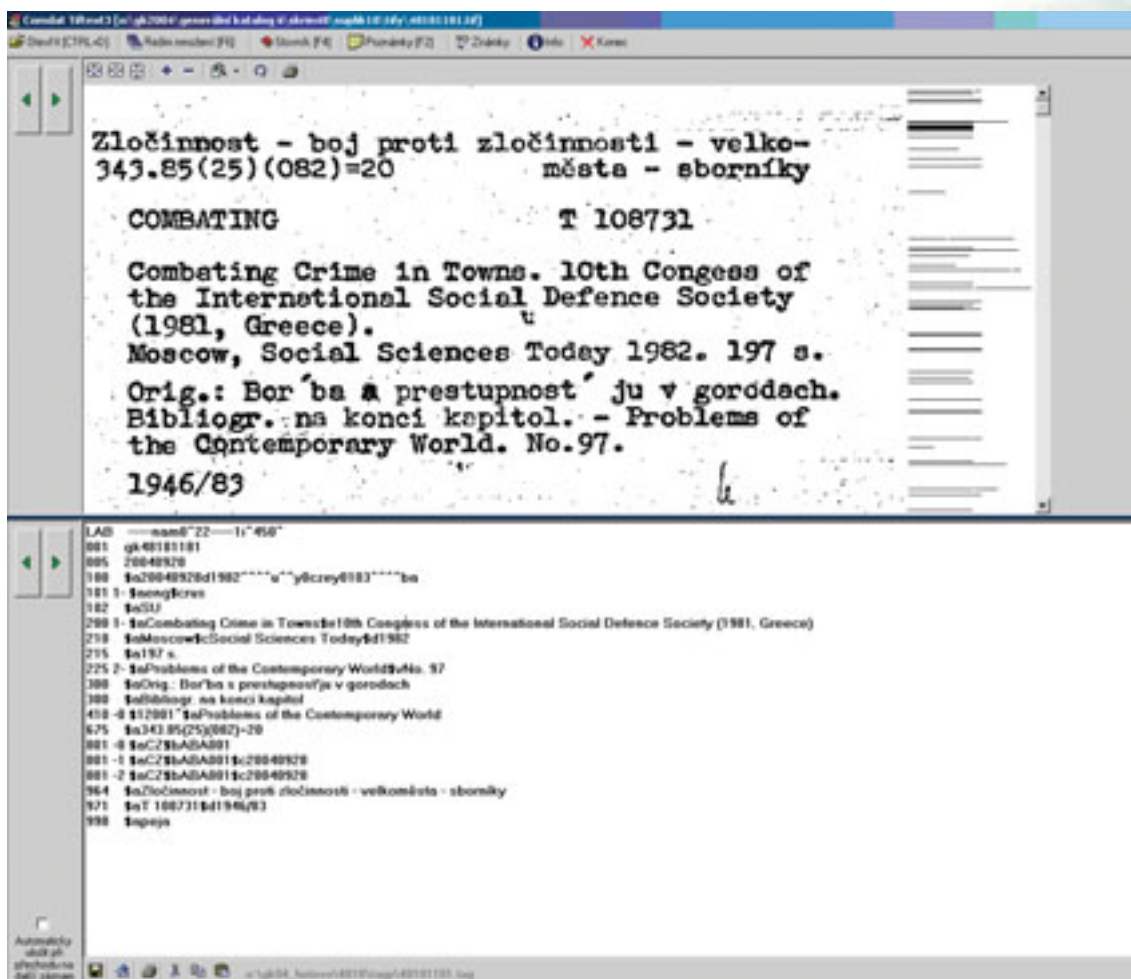
V této etapě dochází ke strukturování textu do podoby odpovídající národním nebo mezinárodním standardům. Pro tento účel byl vyvinut program Comdat ProTag, který umožňuje snadnou a efektivní tvorbu strukturovaných záznamů. Jednotlivá slova nebo řetězce znaků přenáší odborný pracovník do odpovídajícího pole struktury záznamu, přičemž může využít řadu podpůrných nástrojů (slovníky frekventovaných výrazů, automatické strukturování v případech, kde je možné stanovit jednoznačný algoritmus, náhrada velkých a malých písmen a/nebo celých slov atp.) . Obr. 10 je ukázkou strukturování záznamu do polí a podpolí bibliografického formátu UNIMARC. Uvedenou „masku“ je však snadné modifikovat pro libovolný výstupní formát.



Obr. 10

2.4 Revize

Závěrečnou fází retrospektivní konverze je provedení revize zpracovaného záznamu. Revize záznamu je nevyhnutelná pro dosažení potřebné kvality záznamů. Revizi strukturované podoby záznamu obvykle provádí zkušený knihovník na základě původní obrazové podoby záznamu. K provádění revizí slouží program Comdat TifText (viz obr. 11).



Obr. 11

3 RIS – systém pro řízení retrospektivní konverze

Systém RIS jako celek je cenným nástrojem pro správu a řízení retrospektivní konverze zejména ve velkých institucích, kde se jedná o dlouhodobý a několikavrstvý proces a řízení velkých týmů pracovníků. Některé jeho moduly však mohou být využitelné univerzálně.

Systém RIS má třívrstvou architekturu na principu klient-server a je instalován v síťovém prostředí knihovny. Základní vrstvou systému tvoří databázová vrstva. Střední vrstvou je část business logiky poskytující služby všem uživatelům, kteří mají instalovánu třetí, klientskou část systému.

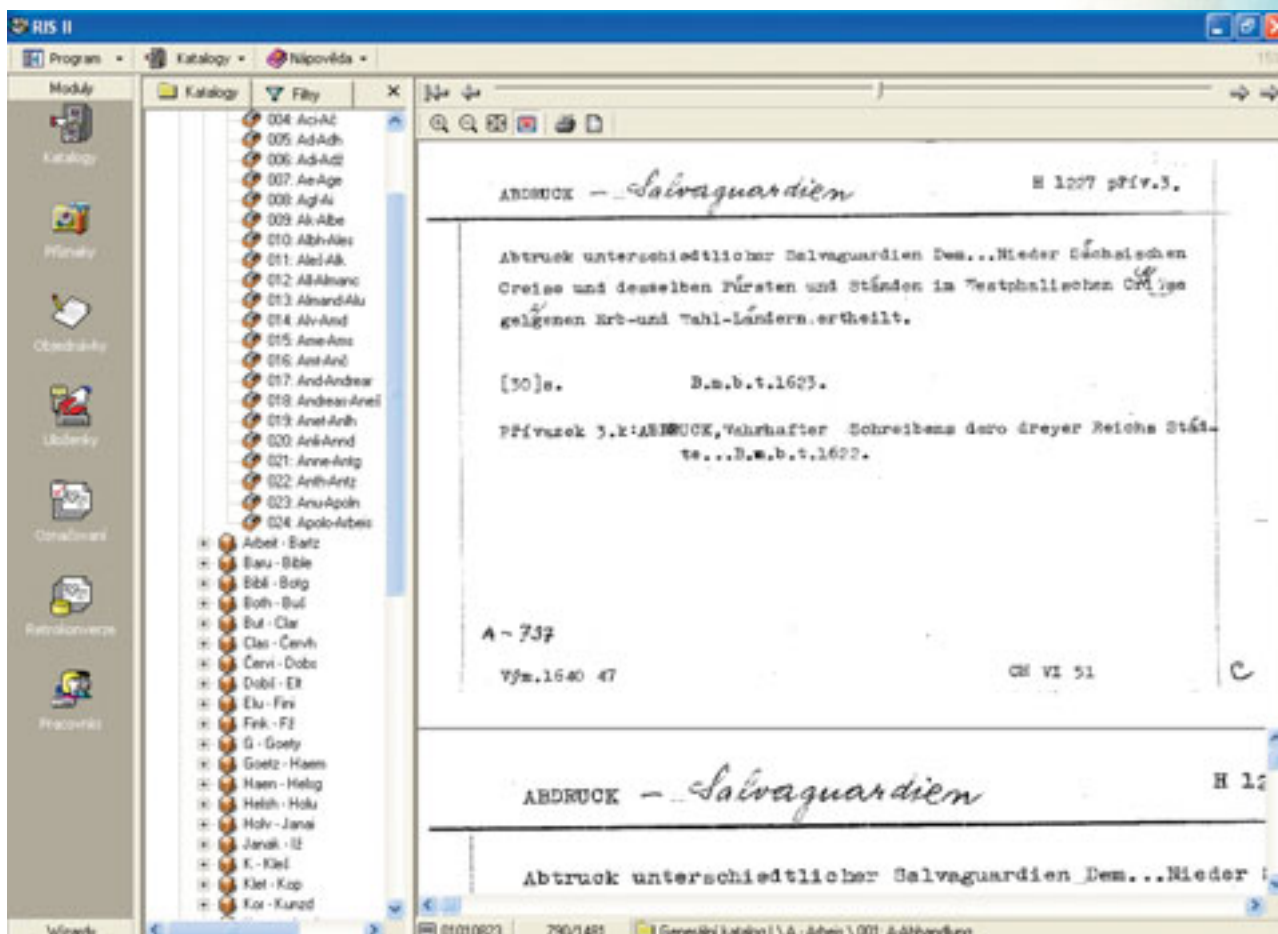
Funkce systému RIS umožňující práci s naskenovanými záznamy jsou rozděleny do pěti základních oblastí:

1. Prohlížení záznamů (modul KATALOGY)
2. Správa objednávek (modul OBJEDNÁVKY)
3. Označování záznamů (modul PŘÍZNAKY)
4. Uložení (modul ULOŽENKY)
5. Řízení retrospektivní konverze (modul RETROKONVERZE)

3.1 Prohlížení záznamů

Prohlížení záznamů je základní funkcí systému a probíhá v modulu KATALOGY, který umožňuje procházet katalogy a zobrazovat záznamy. Zobrazování záznamů je optimalizováno pro potřeby pracovníků knihoven. Uživatel může záznamy pomocí lupy zvětšovat, zmenšovat a používat řadu dalších nástrojů pro zlepšení čitelnosti záznamu (viz obr. 12).

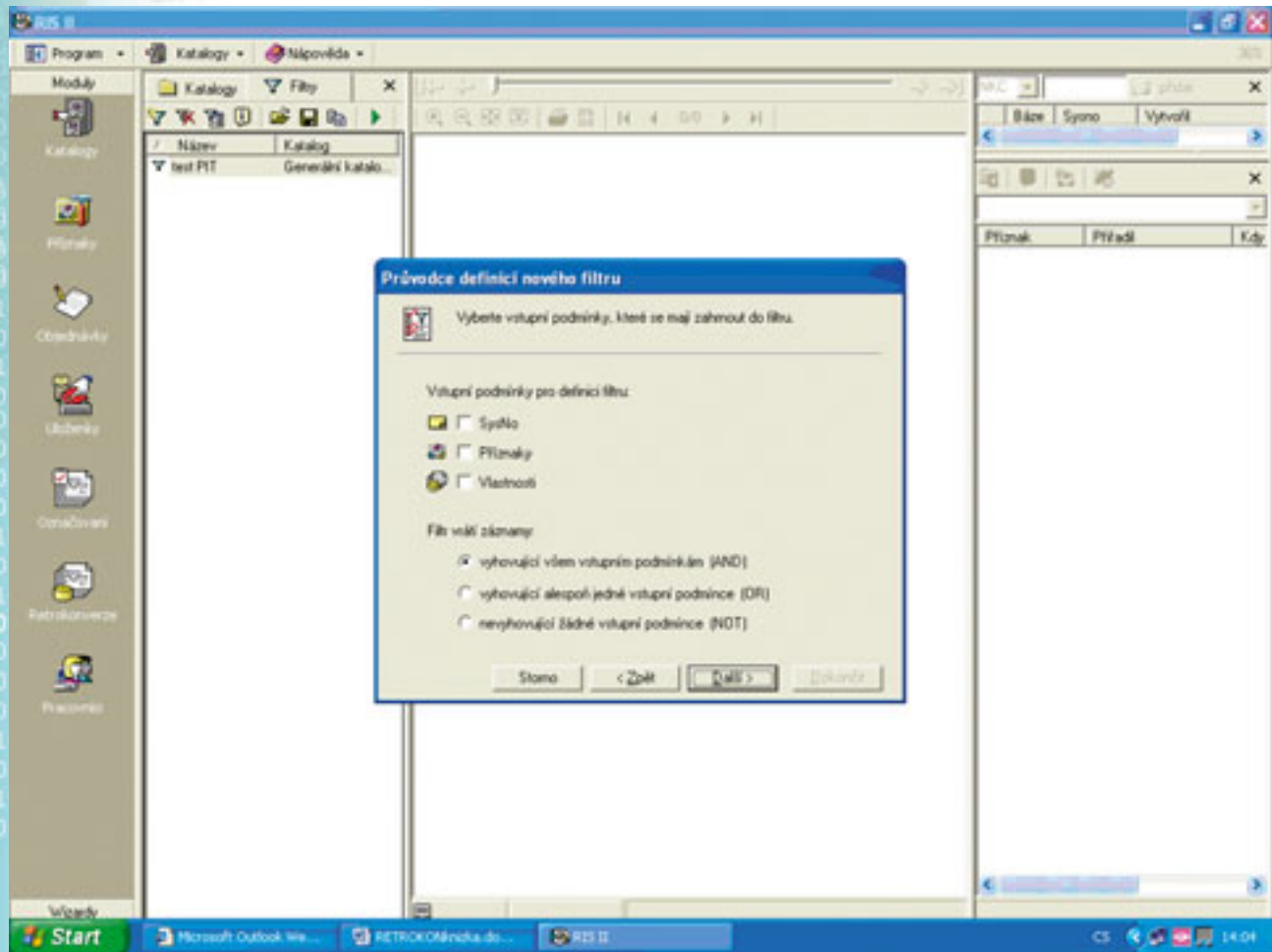
Významným přínosem je sada funkcí pro manipulaci se záznamy. Díky této sadě jsou k dispozici nástroje pro vkládání, přesuny, spojování a rozpojování naskenovaných záznamů. Manipulace se záznamy je významná z toho důvodu, že v naskenovaných, stejně jako v lístkových katalogích mohou být chyby. Chyby mohou vznikat v průběhu využívání lístkového katalogu před jeho skenováním či při procesu skenování. Nástroje pro manipulaci se záznamy umožňují opravovat chyby v katalogu nebo vlastních záznamech operativně při jejich zjištění. Není tedy nutné žádat dodavatelskou firmu nebo správce systému, aby provedl požadovanou



Obr. 12

operaci. Vkládání nově naskenovaných záznamů je stejně snadné jako oprava chyb. Záznamy lze během prohlížení také třídit do skupin na základě předem definovaných filtrů (viz obr. 13). Filtry jsou vytvářeny v návaznosti na příznaky přiřazované jednotlivým záznamům na základě jejich vlastností a umožňují tak třídění celého katalogu podle různých kritérií pro následnou úplnou retrospektivní konverzi.

Příznakům se budeme podrobněji věnovat v kapitole 3.3 Označování záznamů.



Obr. 13

3.2 Správa objednávek

V tomto modulu jsou poskytovány služby související se zpracováním a vyřizováním objednávek (viz obr. 14). Objednávky, které byly vytvořeny čtenáři prostřednictvím www rozhraní na internetu, nebo pracovníky jiných knihoven v rámci meziknihovni výpůjční služby, jsou zpracovávány systémem RIS.

The screenshot shows the RIS II software interface. The window title is "RIS II". The menu bar includes "Program", "Objednávky", and "Nápověda". The left sidebar contains a "Moduly" (Modules) list with icons for "Katalogy", "Přihlasy", "Objednávky", "Uložení", "Cenařování", "Retriekvence", and "Pracovníci". The main area is titled "Hala služeb" (Service Hall) and contains a search bar with the number "122" and a "Přijato" (Received) button. Below the search bar are tabs for "Oslo", "Signatura", and "Přijato". The right pane is titled "Žádná objednávka" (No order) and contains a form with the following fields: "Autor:" (Author), "Název:" (Title), "Signatura:" (Signature) with two input boxes, and "Uložení:" (Storage) with one input box. At the bottom of the right pane is a table with the following columns: "Signatura", "Díl, svazek, roč." (Part, volume, year), "Uložení" (Storage), and "Stav" (Status). The table is currently empty.

Obr. 14

System registruje objednávky a vede evidenci jejich zpracování. Po přijetí a registraci objednávky systém automaticky vytiskne objednávací lístek, který slouží pro interní potřebu knihovny při procesu vyhledání a následném předání čtenáři. Další informace o průběhu zpracování objednávky dostává čtenář průběžně elektronickou poštou nebo si je může na základě čísla objednávky vyhledat přes www rozhraní.

Objednávkový systém vede samostatně evidenci objednávek a statistiku pro každou studovnu zvlášť. Knihovna pak má přehled o počtu objednávek v jednotlivých studovnách a úspěšnosti jejich zpracování (viz obr. 15).

Statistika

Knihovni služba: **Všechny knihovni služby**

		Kladné vy...	Záporné v...	Nevyřízené
[-] Všechny knihovni služby	Slovenská knihovna	0	0	0
[-] 2001	Hala služeb	0	0	0
[-] 2002	Referenční centrum (MVS)	0	0	0
[-] 2003	Všobecná studovna	0	0	0
[-] 2003	Studovna vědeckých pracovníků	0	0	0
[-] 2004	Studovna společenských a přírodních věd	0	0	0
[-] 2004	Slovenská knihovna (MVS)	0	0	0
[+] Únor	5. listopad 2004	0	0	0
[+] březen	6. listopad 2004	0	0	0
[+] duben	7. listopad 2004	0	0	0
[+] květen	8. listopad 2004	0	0	0
[+] červen	9. listopad 2004	0	0	0
[+] červenec	10. listopad 2004	0	0	0
[+] srpen	11. listopad 2004	0	0	0
[+] září	12. listopad 2004	0	0	0
[+] říjen	13. listopad 2004	0	0	0
[+] listopad	14. listopad 2004	0	0	0
[+] prosinec	15. listopad 2004	0	0	0
	16. listopad 2004	0	0	0
	17. listopad 2004	0	0	0
	18. listopad 2004	0	0	0
	19. listopad 2004	0	0	0

Zavřít

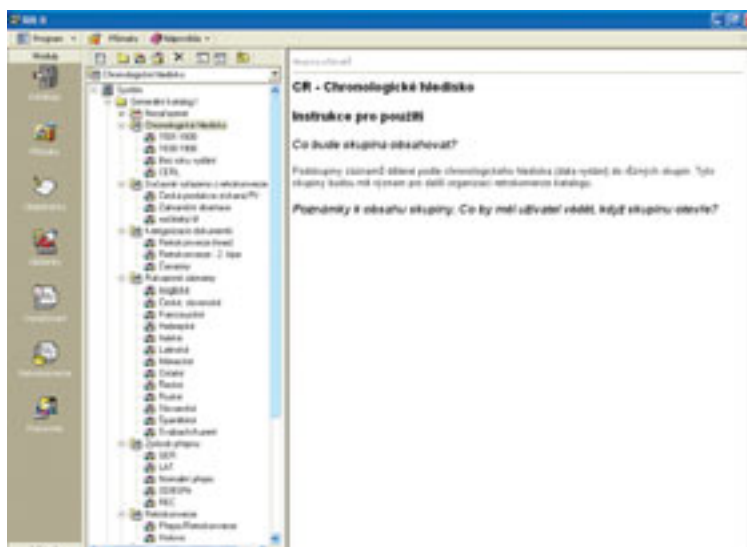
Obr. 15

3.3 Označování záznamů

Označování záznamů je základním nástrojem pro seskupování a třídění záznamů podle předem daných kritérií. Pro označování záznamů slouží příznaky. Příznaky jsou volně definovatelné na základě daných kritérií. Odborný pracovník si předem pro každý katalog nadefinuje skupinu příznaků, která charakterizuje vlastnosti záznamů v určitém katalogu (viz obr. 16). Z této skupiny poté vybírá jednotlivé příznaky a přiřazuje je ke konkrétním záznamům. Příznaky je možné přesouvat mezi katalogy, aby nebylo nutno totožné příznaky definovat pro každý katalog zvlášť.

Příklady kritérií pro definování příznaků:

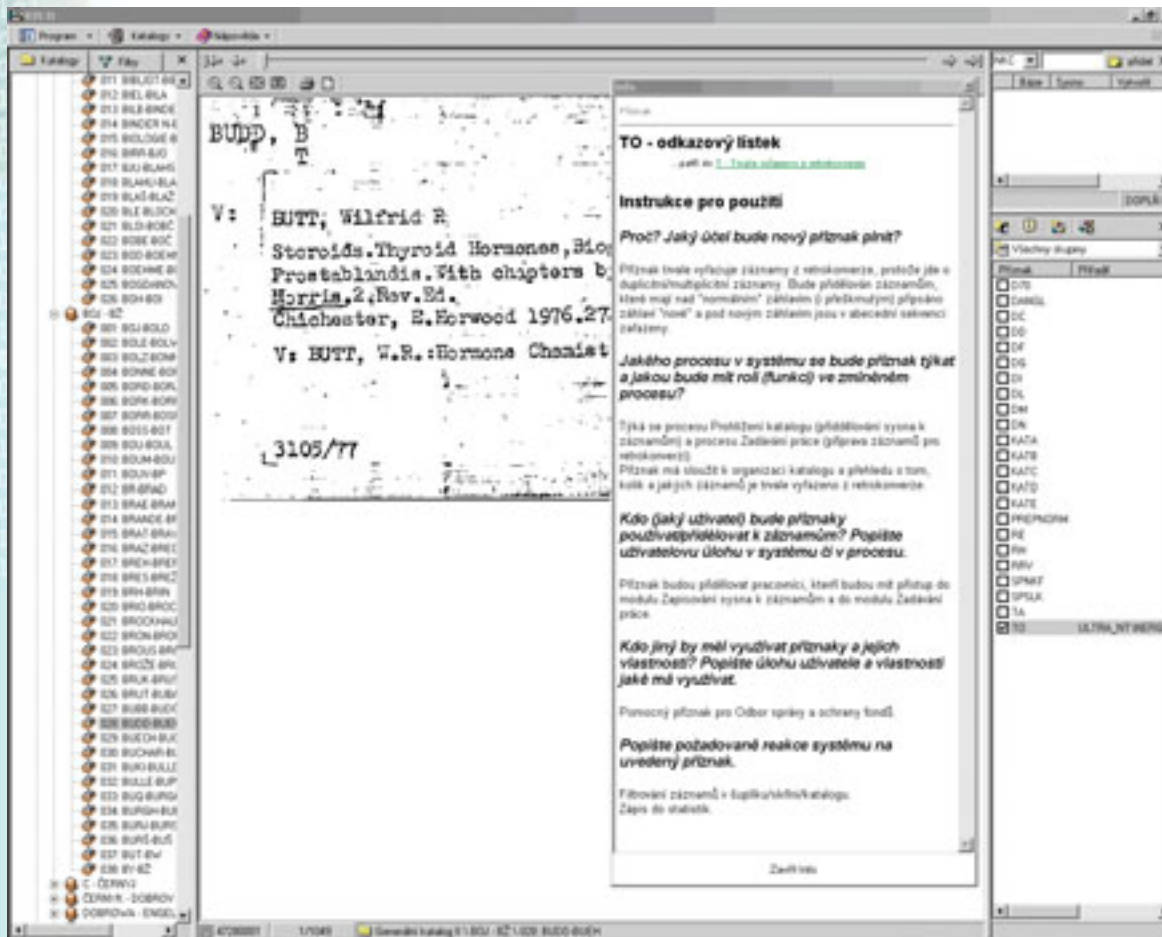
- jazyk
- časové hledisko
- rukopisné záznamy
- specifický jazyk
- specifické písmo
- způsob retrokonverze
- vyřazení z retrokonverze
- ... a další libovolné příznaky dle požadavku uživatele



Obr. 16

Podle příznaků jsou záznamy následně seskupovány a přiřazovány odborným pracovníkům, kteří je v rámci retrospektivní konverze zpracovávají. Případně je s nimi nakládáno jinak v souvislosti s jejich specifickým charakterem.

Při procházení záznamů jsou příznaky zobrazovány ke každému záznamu, s možností vyvolání podrobných informací o jednotlivých příznacích (viz obr. 17). K výše popsanému seskupování jsou využívány filtry, které již byly zmíněny v kapitole 3.1 Prohlížení záznamů.



Obr. 17

3.4 Uložky

Uložky slouží ke zpřístupnění dokumentů, které nejsou běžně přístupné ve skladištích, ačkoliv jsou řádně katalogizovány. Jedná se především o ty dokumenty, které jsou součástí příručních knihoven jednotlivých studoven či odborných útvarů, a dokumenty ztracené, zcizené nebo odepsané. Pro správu záznamů uložek a jejich signatur slouží samostatný modul (viz obr. 18). Tento modul umožňuje vytváření a správu uložek a jejich seznamů, se kterými se následně pracuje při vyřizování objednávek a dalších agend knihovny.

Signatura	Díl svazek, sv.	Uložení	Stav	Znáení	Kdy
1 E 00		Fond 19. století			1.1.2004 0:00
1 F 601		Fond 19. století			1.1.2004 0:00
1 H 136		Fond 19. století			1.1.2004 0:00
1 X 1	4	Comdat	Test	ppučk	26.11.2004 ...
1 X 2	1			SNYDRM	26.4.2004 1...
1 X 3				KATALOG	26.4.2004 1...
-1 X 3	1	BMMS-Knihopis	Ztráta	JERABKOVÁJ	26.4.2004 1...
-1 X 3	2	BMMS-Knihopis	Ztráta	JERABKOVÁJ	26.4.2004 1...
-1 X 3	3	BMMS-Knihopis	Ztráta	KATALOG	26.4.2004 1...
1 X 5	sv.1	OOKF-Hodviva/Přilucka	Ztráta	KATALOG	26.4.2004 1...
-1 X 5	sv.2	BMMA	Ztráta	KATALOG	26.4.2004 1...
-1 X 5	sv.3	OOKF-Hodviva/Přilucka	Revizní ztráta	KATALOG	26.4.2004 1...
1 X 12	sv.1	BMCH/studovna	Revizní ztráta	SNYDRM	26.4.2004 1...
-1 X 12	sv.2	BMCH/studovna	Revizní ztráta	SNYDRM	26.4.2004 1...
-1 X 12	sv.3	BMCH/studovna	Revizní ztráta	SNYDRM	26.4.2004 1...
9 D 1067		Polkozeno			1.1.2004 0:00
9 G 735		Polkozeno			1.1.2004 0:00
9 G 756	Sv.1	Polkozeno			1.1.2004 0:00
9 G 5727		Polkozeno			1.1.2004 0:00
9 G 8186		Polkozeno			1.1.2004 0:00
9 H 792	Bd.16	Revizní ztráta			1.1.2004 0:00
9 J 3674	191860	Polkozeno			1.1.2004 0:00
9 K 1697	Bd.2	Polkozeno			1.1.2004 0:00
9 L 239	191809	Polkozeno			1.1.2004 0:00
12 A 141		Revizní ztráta			1.1.2004 0:00
12 F 2909		Revizní ztráta			1.1.2004 0:00
15 B 299		Revizní ztráta			1.1.2004 0:00
18 D 764		Fond 19. století			1.1.2004 0:00
18 D 2646		Revizní ztráta			1.1.2004 0:00
18 G 1580		Fond 19. století			1.1.2004 0:00
18 J 1514		Fond 19. století			1.1.2004 0:00
19 G 335		Fond 19. století			1.1.2004 0:00
20 B 120	T.9	Revizní ztráta			1.1.2004 0:00
21 F 1199		Revizní ztráta			1.1.2004 0:00
23 D 594		Revizní ztráta			1.1.2004 0:00
21 E 949		Revizní ztráta			1.1.2004 0:00

Obr. 18

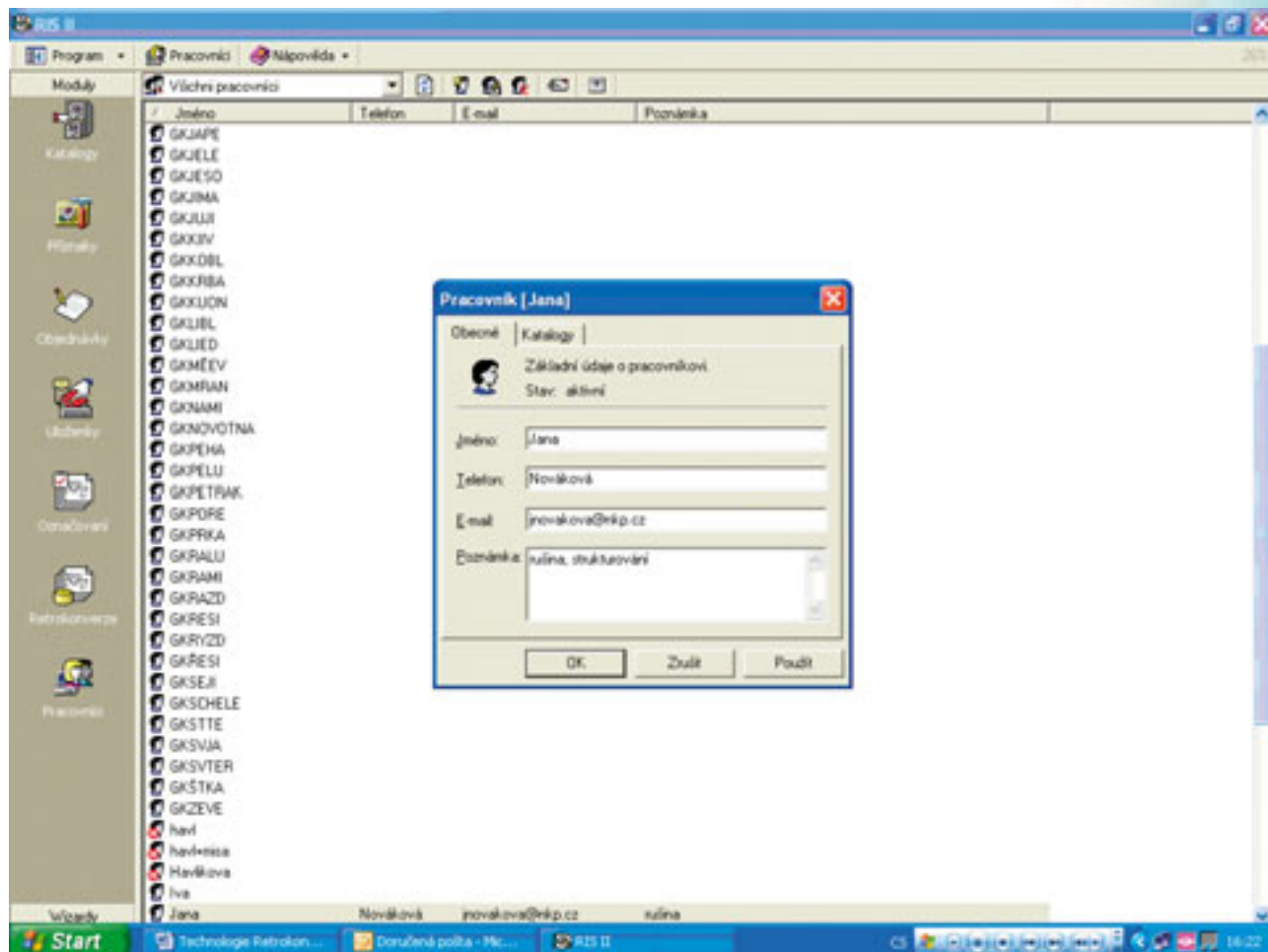
3.5 Řízení retrospektivní konverze

Hlavní službou poskytovanou systémem RIS je řízení retrospektivní konverze. Jak bylo uvedeno v úvodu, technologie RETROKON probíhá ve čtyřech fázích. Tyto fáze na sebe bezprostředně navazují. Výstupem každé z fází je určitá forma záznamu, která je meziproduktem, s nímž se dále pracuje na vstupu fáze následující. Výstupem celého procesu je pak záznam v požadované struktuře a odpovídající kvalitě. V uvedené tabulce jsou znázorněny jednotlivé etapy, procesy, ke kterým v nich dochází, a výstupní tvar záznamu v dané etapě.

	proces	podoba záznamu
I. etapa	skenování	obrazová
II. etapa	přepis	nestrukturovaný text
III. etapa	strukturování	strukturovaný text
IV. etapa	revize	zrevidovaný strukturovaný text

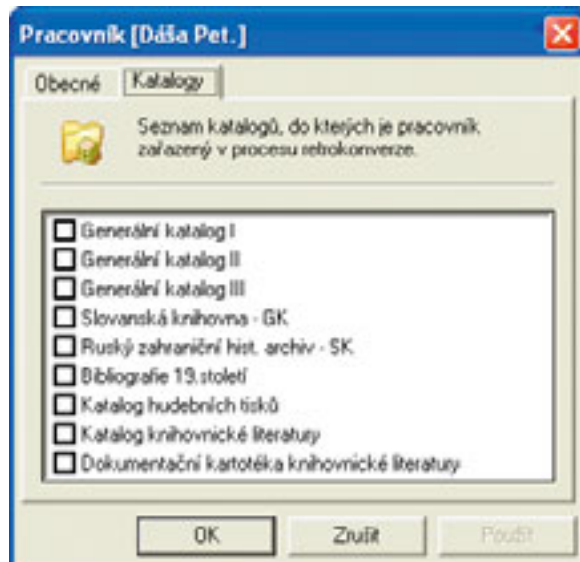
I. etapa tvoří samostatný celek a jejím cílem je vytvoření obrazové podoby katalogu a její zpřístupnění na internetu. Vzhledem k tomu, že se jedná o etapu se specifickým způsobem řízení a zároveň jsou zde často zpracovávány katalogy, u nichž se nepočítá s úplnou retrospektivní konverzí, není zahrnuta do RIS. Proto je RIS navržen pro řízení etap následujících po skenování záznamů.

Pro řízení průběhu retrospektivní konverze je nutné vytvořit nejprve evidenci pracovníků (viz obr. 19), kteří se na retrospektivní konverzi katalogů budou podílet. Pro tyto pracovníky jsou v dalším procesu vytvářeny pracovní dávky (skupiny záznamů), které systém předává dále ke zpracování.



Obr. 19

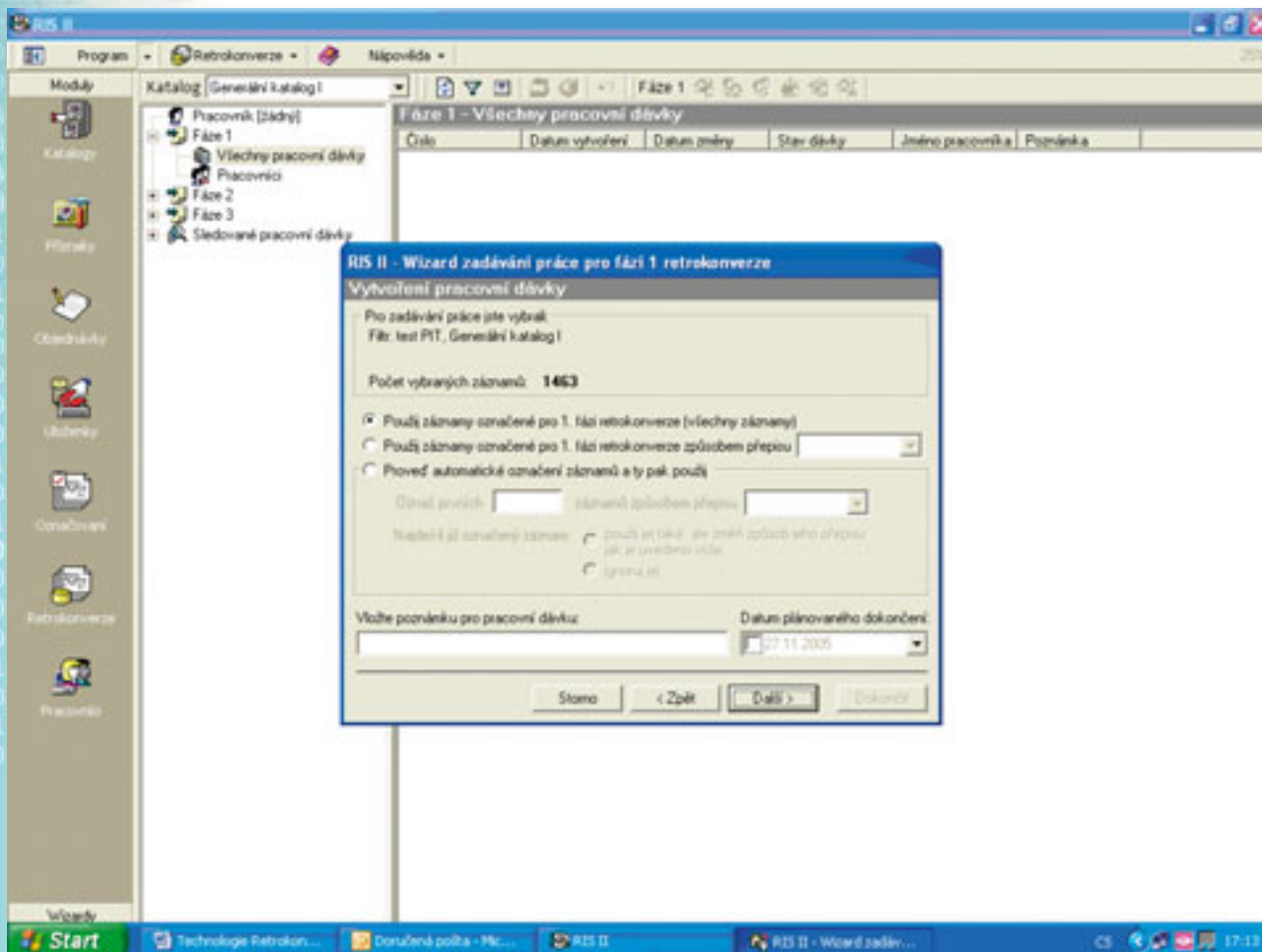
Každému pracovníkovi jsou s ohledem na jeho kvalifikaci přiřazeny katalogy, na jejichž zpracování se bude podílet (viz obr. 20).



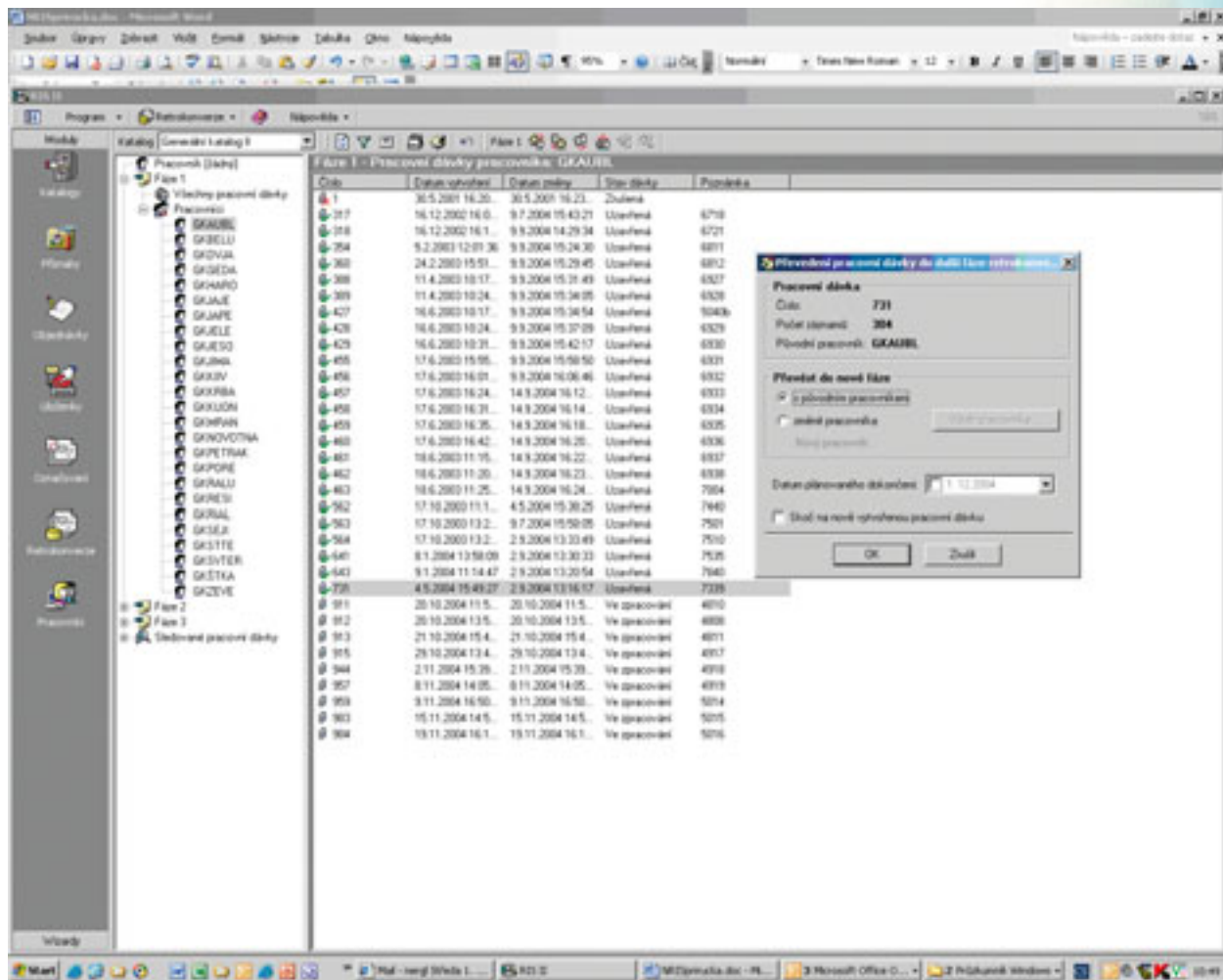
Obr. 20

První etapou retrospektivní konverze, kterou systém RIS eviduje, je přepis. V této etapě jsou záznamy převáděny z naskenované obrazové podoby do podoby textové. V této fázi dochází k vytvoření pracovní dávky pro konkrétního pracovníka. Záznamy lze pro dávku vybírat dvěma způsoby. Buď je možné provést výběr celé zásuvky z katalogu, nebo je možné vytvořit dávku na základě příznaků, s využitím dříve připravených filtrů (viz obr. 21).

Vybrané záznamy jsou následně exportovány ke zpracování (viz obr. 22).



Obr. 22



Obr. 23

Pracovník provede zpracování záznamu nezávisle na systému (v tomto případě se jedná o převod záznamů do textového tvaru pomocí programu TextTif) a následně jej již v nové podobě vrátí zpět. Takto zpracované záznamy jsou uloženy do systému RIS a připraveny k exportu dalšímu pracovníkovi, který bude zajišťovat následující etapu – převod záznamů do strukturované podoby. Jednotlivé dávky je nutné uzavřít a následně je možný jejich přesun do další fáze retrospektivní konverze (viz obr. 23).

Proces strukturování záznamů probíhá ve speciálním programu ProTag. Na strukturování záznamů navazuje proces revize (v programu TifText). V rámci revize je provedena závěrečná kontrola záznamů před exportem do knihovního systému, což je poslední krok retrospektivní konverze metodou RETROKON.

Jednotlivé etapy zpracování jsou na sobě nezávislé a pracovníci podílející se na zpracování se mohou podle své kvalifikace zapojit do jedné nebo více etap. V průběhu zpracování jsou všechny podoby záznamu ukládány a archivovány. To znamená, že je možné kdykoliv se vrátit do výchozí pozice kterékoli etapy, aniž by došlo ke ztrátě dat.

Informace o průběhu a výsledku zpracování jsou poskytovány systému NRIS, který je následně publikuje u příslušného katalogu na internetu. Systém RIS je navržen jako modulární. Není tedy třeba využívat a instalovat všechny moduly, pokud by nebyly v určité knihovně využity.

Systém je navržen na platformě serveru MS Windows a databáze MS SQL. Kapacita systému z hlediska počtu záznamů není omezena a je ovlivněna pouze výkonem serveru. Maximální počet záznamů, které tak mohou být v systému, se pohybuje řádově v desítkách milionů. Správa, administrace a práce s RIS je řízena na principu přístupových práv řízených prostřednictvím uživatelských účtů.

4 Závěr

V současné době existuje v České republice kvalitní, prakticky dokonale ověřený a spolehlivý nástroj pro retrospektivní konverzi katalogů, který byl využit v řadě našich knihoven a s určitými modifikacemi i v knihovnách zahraničních. Do Národní knihovny ČR přijíždí každoročně řada zahraničních návštěv, které se zajímají právě o technologii RETROKON. Nejedná se jen o pracovníky knihoven, ale stále častěji i o pracovníky archivů, muzeí, galerií a dalších institucí. Technologie RETROKON může být snadno modifikována i pro jejich potřeby a mohou ji využít buď jako celek, nebo její jednotlivé moduly. Významným obohacením technologie bude podpora protokolu Z39.50, která umožní využití bibliografických i autoritních záznamů z externích zdrojů.

Cílem této publikace bylo představit technologii RETROKON v celé její komplexnosti pracovníkům knihoven, kteří většinou znají pouze některé z modulů, dále pak pracovníkům archivů, muzeí, galerií a dalších paměťových institucí, jež hledají vhodný způsob pro převod svých rozsáhlých katalogů, kartoték a nejrůznějších soupisů do digitální podoby a jejich následné zpřístupnění prostřednictvím internetu. Věříme, že technologie RETROKON pro ně může být řešením nebo alespoň inspirací.

5 Další informace a kontakty

Další užitečné informace a dokumenty naleznete na těchto adresách:

<http://retrokon.nkp.cz>

<http://visk.nkp.cz/VISK5.htm>

Na vaše dotazy a podněty se těší řešitelský tým RETROKON.

Knihovnické dotazy k technologii RETROKON zasílejte na adresu bohdana.stoklasov@nkp.cz,
technické dotazy na adresu mbares@comdat.cz.





1101101000
1001101011

1001010010 011
1010101101 111
1010010001 111

RETROKON

Technologie RETROKON jako komplexní nástroj pro digitalizaci a zpřístupnění katalogů, kartoték a soupisů prostřednictvím internetu

Bohdana Stoklasová, Miroslav Bareš, Anna Nerglová

Vydavatel: Národní knihovna České republiky, 2004

Grafický návrh: Martin Zhouf

DTP studio: POLYSERVIS, Chrustenice

Tisk: DECIBEL PRODUCTION s.r.o.

Distribuce: Národní knihovna České republiky
oddělení odbytu
Štěrboholská 55, 102 00 Praha 15
Tel: 420 281 013 230
E-mail: mirosovsky.ivo@cdh.nkp.cz

