



Projekt CDArcha: kooperativní zpracování datových CD disků

Zdeněk Hruška, Petr Žabička

Moravská zemská knihovna v Brně

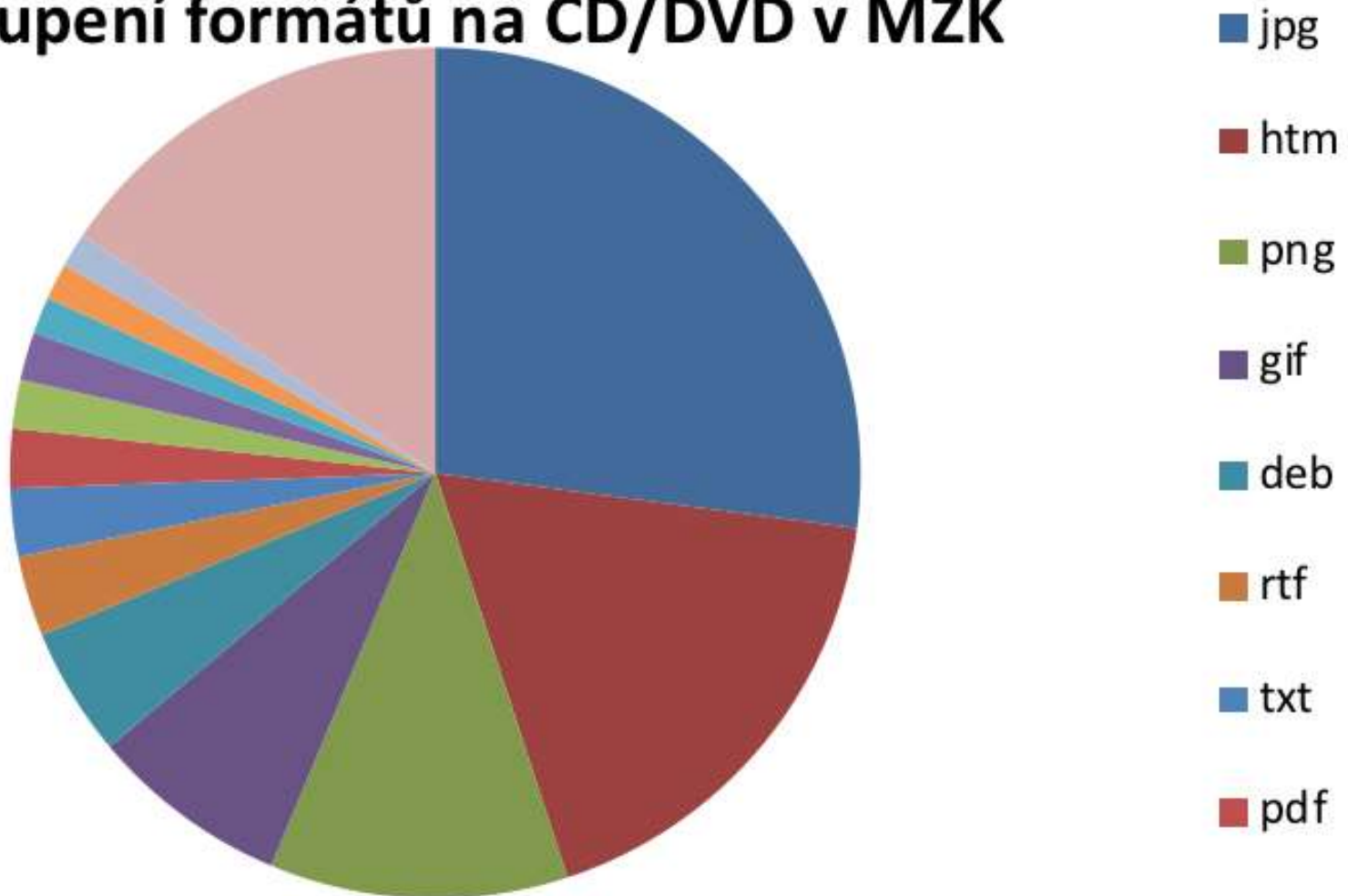
AKM 2018

- ❑ +/- 25 000 CD/DVD disků v MZK
- ❑ Nejstarší disky z 90. let
- ❑ Hrozí degradace datové vrstvy, poškrábání, ztráta, krádež, ...
- ❑ Otázka zpřístupnění uživatelům a dlouhodobá ochrana dat

- ❑ **Samostatná CD/DVD (povinný výtisk)**
- ❑ **Samostatný záznam**
 - ❑ Skripta, sborníky z konferencí, výstupy projektů, ...
- ❑ Velké množství formátů: .doc, .ppt, .xls, .pdf, .jpg, .html a další...

- ❑ **Příloha** u monografií a periodik
- ❑ **Záznam mateřského dokumentu**
 - ❑ Datová CD/DVD
 - ❑ Video DVD, video CD
 - ❑ SW, soubory, multimédia, ...
 - ❑ (Zvuková CD)
 - ❑ (Multisession CD (zvuk + data))

Zastoupení formátů na CD/DVD v MZK



Samostatná CD/DVD,
formáty:

Formát	Počet
jpg	54083
htm	35755
png	22649
gif	15150
deb	9788
rtf	6104
txt	5163
pdf	4489
bmp	3699
html	3645

- ❑ Samostatná CD “archivujeme” od roku 2013 - zpracováno cca 1800 disků
- ❑ Záloha na externí disk
- ❑ Konverze na PDF
 - ❑ Formáty MS Office + obrázky
 - ❑ **Import do Krameria**
 - ❑ Pokud je celý obsah v Krameriu, disk už se nepůjčuje.

- ❑ 2018, projekt “Vytvoření nástrojů pro kooperativní zpracování a bit-level ochranu obsahu datových optických disků ve fondu knihoven” (VISK3)
- ❑ Připraveno na možnost spolupráce knihoven při zpracování fondu datových CD/DVD.

V rámci projektu vzniká:

- ❑ Návrh metodiky pro zpracování a bit-level ochranu obsahu datových CD/DVD
- ❑ Návrh standardu datového balíčku pro obsah datových CD/DVD

- ❑ SW na extrakci dat z CD/DVD
(modifikace SW Obálky knih)
- ❑ **CDArcha-client**
 - ❑ Desktopová aplikace
 - ❑ Metadata z katalogu
 - ❑ Vytvoření ISO image disku
 - ❑ Kontrolní součet (md5)
 - ❑ Sken disku + přebalu + bookletu

- ❑ CDArcha-server
 - ❑ Serverová aplikace
 - ❑ DROID - identifikace formátů
 - ❑ Dokončení balíčku (SIP)
 - ❑ Bit-level ochrana

Hlavní složka obsahuje:

složku **data** + soubory *manifest-md5.txt* a *bagit.txt* (jedná se tedy o BagIt)
podsložky ve složce **data**:

- ☐ **isoimage** - obsahuje jeden či více iso obrazů
- ☐ **mastercopyscan** - obsahuje skeny disku, bookletu, obalu, ...
- ☐ **usercopyscan** - skeny disku, bookletu, obalu, ..., uživ. kopie
- ☐ **amdsec** - obsahuje administrativní metadata
- ☐ **alto** - obsahuje alto z OCR
- ☐ **txt** - obsahuje OCR ve formátu txt

soubory ve složce **data**:

- ☐ *droid_uuid.xml* - výstup z identifikace formátů pomocí DROIDu
- ☐ *info_uuid.xml* - popis obsahu balíčku + seznam souborů
- ☐ *main_mets_uuid.xml* - hlavní METS balíčku
- ☐ *md5_uuid.xml* - kontrolní součet

CDArcha-client - zpracování média



Ve vašem dočasném adresáři jsou archivované média.

[Zobrazit složku](#)

Nová archivace

Vaše současná verze: 0.1
Podpora: Podporováno

Verze je aktuální

Nejnovější verze: 0.1
Datum: 01/20/2018

[Online nápověda a seznam změn](#)

Založit novou jednotku

Monografie **Periodikum**

Založit novou monografii

Čárový kód

Založit

Zpracování média - metadata



Soubor Nastavení Pomoc

Metadata Zpracování Skenování

Informace o části

Název

Juniorstav ... : ... odborná konference doktorského

Autoři

Rok vydání

1999

Odpojit souborný záznam

Číslo části

9/2006

Vyplňte
číslo a
název části

Název části

Juniorstav 2006

Identifikátory SZ

ISBN / EAN / UPC

80-214-2560-1

ČNB

cnb002176734

OCLC

(OCoLC)724299817

ISMN

URN:NBN

Vlastní

Zobrazit metadatový záznam

Metadatový záznam

Plný metadatový záznam

<<< Předchozí záznam 1/1 Následující záznam >>>

SYSNO nkc20112176734
LDR 00574cas a2201213 i 4500
001 nkc20112176734
003 CZ PrNK
005 20171031140750.0
007 ta
008 110323c19999999xr ar 1 b0cze
015 ja cnb002176734 jz cnb001704545 jz cnb001790766 jz cnb001850942 jz cnb002029824 jz cnb001298924
035 ja (OCoLC)724299817
040 ja ABA001 jb cze jd OLA001 jd BOA001 je rda
0411 ja cze ja eng ja slo jh cze jh slo
072 7 ja 624 jx Stavební inženýrství j2 Konspekt j9 19
080 ja 69 j2 MRF
080 ja 624 j2 MRF
080 ja (062.534) j2 MRF
080 ja (048.2) j2 MRF
1112 ja Juniorstav (konference) j7 olak2006319348 j4 aut
24510 ja Juniorstav ... : jb ... odborná konference doktorského studia = ... professional conference of postgraduate students : sborník anotací : Vys
2461 ji Název v tiráži: ja Sborník anotací konference Juniorstav ...
2463 ja Odborná konference doktorského studia - Juniorstav
2463 ja Professional conference of postgraduate students - Juniorstav
2463 ja Professional conference of PhD students - Juniorstav
2463 ja International conference of PhD students - Juniorstav

Použít

ihovna

Zpracování média - vložit médium



▼ Doplnkové informace

▲ Krok 1: Existující archivy médií k tomuto bibliografickému záznamu

Obnovit seznam

Zobrazit obsah

Stav archivu: **Ještě neexistuje**
Archiv neexistuje a bude vytvořen s první archivací

Uzavřít archiv

Skenovat

▲ Krok 2: Načtení média z mechaniky

Mechanika:

Název média: **není vloženo**

Císlo svazku:

Velikost:

▲ Krok 3: Archivace

Pořadové číslo disku:

 **Médium není načtené**

Zpracování média - ověření duplicity



▼ Doplnkové informace

▲ Krok 1: Existující archivy médií k tomuto bibliografickému záznamu

Obnovit seznam

Zobrazit obsah

Stav archivu: **Ještě neexistuje**

Archiv neexistuje a bude vytvořen s první archivací

Uzavřít archiv

Skenovat

▲ Krok 2: Načtení média z mechaniky

Mechanika:

Název média: **juniorstav2006**

Číslo svazku: 1C9A6AFD

Velikost: 504,50 MB

▲ Krok 3: Archivace

Pořadové číslo disku:

Médium ještě nebylo archivováno.

Zpracování média - vytváření ISO obrazu I/I

The screenshot displays a software interface for media processing. At the top, a list of five steps is shown, each with a green checkmark icon:

- 1) Příprava (Zasílání identifikátorů)
- 2) Stahování (Kopírování média na lokální PC)
- 3) Kontrola (Výpočet kontrolního součtu média)
- 4) Zaslání (Kopírování média na server)
- 5) Kontrola (Kontrola zasláných dat a ukončení)

Below the list is a horizontal progress bar. The first 75% of the bar is green, and the remaining 25% is grey. To the right of the bar, the text "75%" is displayed.

Below the progress bar, the following text is shown:

Uplynulý čas: 00:00:00
Zkopírování: 382 MB

Below this text is a log window titled "Log procesu archivace". The log contains the following entries:

- 11.09.2018 11:40:49 Kompletace identifikátorů záznamů.
- 11.09.2018 11:41:01 Chystání metadatových souborů MARC XML a MODS.
- 11.09.2018 11:41:01 Zaslání identifikátorů a metadat.
- 11.09.2018 11:41:01 Odeslání identifikátorů záznamu a metadat úspěšně.
- 11.09.2018 11:41:01 Vytváření ISO na lokální PC.

At the bottom of the interface, there is a section titled "Skenování" with a dropdown arrow.

Zpracování média - kopírování na server I/I

The screenshot displays a software interface for media processing. At the top, a list of five steps is shown, each with a green checkmark icon:

- 1) Příprava (Zasílání identifikátorů)
- 2) Stahování (Kopírování média na lokální PC)
- 3) Kontrola (Výpočet kontrolního součtu média)
- 4) Zaslání (Kopírování média na server)
- 5) Kontrola (Kontrola zaslanych dat a ukončení)

Below the list is a green progress bar that is approximately 65% full. To the right of the bar, the text "65%" is displayed. Below the progress bar, the following information is shown:

Uplynulý čas: 00:00:00
Zkopírováno: 327 MB

Below this information is a log window titled "Log procesu archivace". The log contains the following entries:

- 11.09.2018 11:40:49 Kompletace identifikátorů záznamů.
- 11.09.2018 11:41:01 Chystání metadatových souborů MARC XML a MODS.
- 11.09.2018 11:41:01 Zaslání identifikátorů a metadat.
- 11.09.2018 11:41:01 Odeslání identifikátorů záznamu a metadat úspěšně.
- 11.09.2018 11:41:01 Vytváření ISO na lokální PC.
- 11.09.2018 11:46:10 Ukládání ISO úspěšně.
- 11.09.2018 11:46:13 Kontrolní součet lokálního ISO souboru: 69f5e8a9959c8ce2b96417eff7fa2c31
- 11.09.2018 11:46:13 Odesílám ISO soubor na server.
- 11.09.2018 11:46:13 Výpočet kontrolního součtu ukončen.

At the bottom of the interface, there is a tab labeled "Skenování".

Zpracování média - kontrola na serveru



- ✓ 1) Příprava (Zasílání identifikátorů)
- ✓ 2) Stahování (Kopírování média na lokální PC)
- ✓ 3) Kontrola (Výpočet kontrolního součtu média)
- ✓ 4) Zaslání (Kopírování média na server)
- ✓ 5) Kontrola (Kontrola zasláných dat a ukončení)

Uplynulý čas: 00:00:00
Zkopírováno: 503 MB

100%

Log procesu archivace

11.09.2018 13:41:14 Ukládání ISO úspěšné.
11.09.2018 13:41:16 Kontrolní součet lokálního ISO souboru: 69f5e8a9959c8ce2b96417eff7fa2c31
11.09.2018 13:41:16 Odesílám ISO soubor na server.
11.09.2018 13:41:16 Výpočet kontrolního součtu ukončen.
11.09.2018 13:41:29 Upload ISO ukončen.
11.09.2018 13:41:29 Čekám na výpočet kontrolního souboru na serveru max 60s.
11.09.2018 13:41:29 Čekám na dokončení na straně serveru: 59s
11.09.2018 13:41:30 Kontrolní součet vypočtený serverem: 69f5e8a9959c8ce2b96417eff7fa2c31

11.09.2018 13:41:30 **Proces archivace úspěšně ukončen**
11.09.2018 13:41:30 Chystání metadatových souborů MARC XML a MODS.
11.09.2018 13:41:30 Zaslání identifikátorů a metadat.

Skenování



SKENOVAT

Opakovaně načíst obrázek

Zpracování média - skenování



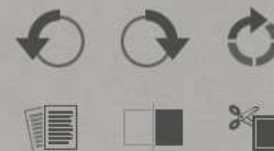
Médium



Booklet
0 stran



Transformace



Výsekový režim



Skenovat médium

A5 A4 A3

Skenovat booklet

A5 A4 A3

Načíst ze soubor

Odeslat

Zpracování média - úspěšně odesláno



- ✓ 1) Příprava (Zaslaní identifikátorů)
- ✓ 2) Stahování (Kopírování média na lokální PC)
- ✓ 3) Kontrola (Výpočet kontrolního součtu média)
- ✓ 4) Zasílání (Kopírování média na server)
- ✓ 5) Kontrola (Kontrola zaslaných dat a ukončení)

Uplynulý čas: 00:00:00

Zkopírováno: 503 MB

Log procesu archivace

11.09.2018 13:41:14 Ukládání ISO úspěšné.
11.09.2018 13:41:16 Kontrolní součet lokálního ISO souboru: 69f5e8a9959c8ce2b96417eff7fa2c31
11.09.2018 13:41:16 Odesílám ISO soubor na server.
11.09.2018 13:41:16 Výpočet kontrolního součtu ukončen.
11.09.2018 13:41:29 Upload ISO ukončen.
11.09.2018 13:41:29 Čekám na výpočet kontrolního souboru na severu max 60s.
11.09.2018 13:41:29 Čekám na dokončení na straně serveru: 59s
11.09.2018 13:41:30 Kontrolní součet vypočtený serverem: 69f5e8a9959c8ce2b96417eff7fa2c31

11.09.2018 13:41:30 Proces archivace úspěšně ukončen.
11.09.2018 13:41:30 Chystání metadatových souborů MARC XML a MODS.
11.09.2018 13:41:30 Zasílání identifikátorů a metadat.
11.09.2018 15:15:59 Kompletace dat pro odesílání obálek a bookletů.

Skenování

Úspěšně odesláno



Zpracování média - další!



Metadata

Zpracování

Skenování

Název Juniorstav ... : ... odborná l

Autoři

Rok vydání 1999?-

Identifikátory

Vlastní BOA001-nkc20112176734

Archivovat další médium stejného titulu (bib. záznamu).

Archivovat další médium

Ukončit archivaci současného titulu a vyhledat další titul k archivaci.

Archivace nového titulu

✓ 1) Příprava (Zaslaní identifikátorů)

✓ 2) Stahování (Kopírování média na lokální PC)

✓ 3) Kontrola (Výpočet kontrolního součtu média)

✓ 4) Zasílání (Kopírování média na server)

✓ 5) Kontrola (Kontrola zaslaných dat a ukončení)

Uplynulý čas: 00:00:00

Zkopírováno: 503 MB

Log procesu archivace

11.09.2018 13:41:14 Ukládání ISO úspěšné.

11.09.2018 13:41:16 Kontrolní součet lokálního ISO souboru

11.09.2018 13:41:16 Odesílám ISO soubor na server.

11.09.2018 13:41:16 Výpočet kontrolního součtu ukončen

11.09.2018 13:41:29 Upload ISO ukončen.

11.09.2018 13:41:29 Čekám na výpočet kontrolního součtu

11.09.2018 13:41:29 Čekám na dokončení na straně serveru

11.09.2018 13:41:30 Kontrolní součet vypočtený serverem

11.09.2018 13:41:30 Proces archivace úspěšně ukončen.

11.09.2018 13:41:30 Chystání metadatových souborů

11.09.2018 13:41:30 Zasílání identifikátorů a metadat

11.09.2018 13:45:59 Kompletace dat pro odesílání obálky

Skenování

Přehled médií na serveru



CDArcha

BIBLIOGRAFICKÉ ZÁZNAMY

ARCHIV

UŽIVATELÉ

Archiv médií

Filtrovat:

Rozdělané na klientovi

Rozdělané na serveru

Připravené na archivaci

Archivované

Vše

Search



ID archivu	Stav	Vytvořeno	Poslední úprava	Operace
5b1dcae94df2aa1c9c5b184e	Rozdělané na klientovi	11.5.2018 3:5	11.5.2018 3:5	✎ ✕ →
5b1e46f74df2aa1c9c5b1852	Rozdělané na klientovi	11.5.2018 11:55	11.5.2018 11:55	✎ ✕ →
5b1e478b4df2aa1c9c5b1855	Rozdělané na klientovi	11.5.2018 11:57	11.5.2018 11:57	✎ ✕ →
5b1e481b4df2aa1c9c5b1858	Rozdělané na klientovi	11.5.2018 11:59	11.5.2018 11:59	✎ ✕ →
5b1e48a34df2aa1c9c5b185b	Rozdělané na klientovi	11.5.2018 12:2	11.5.2018 12:2	✎ ✕ →
5b1e48fc4df2aa1c9c5b185e	Rozdělané na klientovi	11.5.2018 12:3	11.5.2018 12:3	✎ ✕ →
5b3236044df2aa1c9c5b1864	Rozdělané na klientovi	26.5.2018 14:48	26.5.2018 14:48	✎ ✕ →
5b4319dd4df2aa1c9c5b1867	Rozdělané na serveru	9.6.2018 10:16	9.6.2018 10:16	✎ ✕ →
5b4f30324df2aa1c9c5b186a	Rozdělané na klientovi	18.6.2018 14:18	18.6.2018 14:18	✎ ✕ →
5b6010a39ee24011e33d4613	Rozdělané na klientovi	31.6.2018 9:32	31.6.2018 9:32	✎ ✕ →

Přehled titulů na serveru



CDArcha

BIBLIOGRAFICKÉ ZÁZNAMY

ARCHIV

UŽIVATELÉ

Bibliografické záznamy



ID biblio. záznamu	Titul	Authors	ISBN / ISSN / EAN	NBN	OCLC	Vytvořeno	Poslední úprava	Operace
5b1e49914df2aa1c9c5b1860	First choice : A1	John Stevens, Angela Lloyd, Alena Telínová, Marion Karg, Frauke Mönkeberg, Jana Fenclová, Tomáš Hausner, Maggie Bouqdib, Richard Dawton	-	cnb001746812	(OCoLC)189427947	11.5.2018 12:6	11.5.2018 12:6	→
5b4319dd4df2aa1c9c5b1866	Inventarizace : praktický přůvodce	Jaroslava Svobodová, Josef Podhorský	-	cnb001368110	(OCoLC)51185271	9.6.2018 10:16	9.6.2018 10:16	→
5b6010a39ee24011e33d4612	Krtčí výlety	Jitka Petrová, Jan Hlaváč	80- 87022- 05-X	cnb001699526	(OCoLC)124089332	31.6.2018 9:32	31.6.2018 9:32	→
	Na Pohádku máje : analýza a	Klára Hanáková, Zdeněk Pospíšil	978- 80-	cnb002173412	(OCoLC)724298796			→

Informace o jednom titulu



CDArcha

BIBLIOGRAFICKÉ ZÁZNAMY

ARCHIV

UŽIVATELÉ

Archiv

Bibliografického záznamu: Juniorstav ... : ... odborná konference doktorského studia = ... professional conference of postgraduate students : sborník anotací : Vysoké učení technické v Brně, Fakulta stavební ... /

← Návrat



ID archivu	Stav	Vytvořeno	Poslední úprava	Operace
5b978db08180333e48364968	Rozdělané na serveru	11.8.2018 11:41	11.8.2018 11:41	

Showing 1 to 1 of 1 rows

Obsah Archivu



CDArcha

BIBLIOGRAFICKÉ ZÁZNAMY

ARCHIV

UŽIVATELÉ

Obsah archivu

Archivu s identifikátorem 5b978db08180333e48364968 / vytvořeného 11.8.2018 11:41

[← Návrat](#)



Č.média	Typ	Velikost	Problém s čtením	Opakovaný upload	Operace
1	iso	503.3 MB	Ne	Ne	↓

Showing 1 to 1 of 1 rows

Další připojené soubory



Typ souboru	Velikost souboru	Operace
dmdsec-archive-marc		↓
dmdsec-archive-mods		↓



Archiv

Bibliografického záznamu: Juniorstav ... : ... odborná
Fakulta stavební ... /

[← Návrat](#)

Editace archivu



Stav archivu:

- Ukončené na klientovi ▾
- Rozpracované na klientovi
- Ukončené na klientovi**
- Rozpracované na serveru
- Připravené na archivaci
- Archivované

Uložit

Zavřít

ID archivu

Stav

Vytvořeno

Poslední úprava

Operace

5b978db08180333e48364968

Rozdělané na serveru

11.8.2018 11:41

11.8.2018 11:41

Showing 1 to 1 of 1 rows

- ❑ certifikace metodiky
- ❑ problematika logické ochrany
- ❑ právní rámec kooperace
- ❑ řešení procesu zpřístupnění

Děkuji za pozornost!

Pište na: hruska@mzk.cz

Moravská zemská knihovna v Brně
www.mzk.cz