

---

# **eDeposit AMQP Kramerius Documentation**

*Release 0.1.0*

**Jan Stavel**

September 10, 2015



<b>1</b>	<b>Požadavky</b>	<b>1</b>
<b>2</b>	<b>Omezení</b>	<b>3</b>
<b>3</b>	<b>Export do Krameria</b>	<b>5</b>
3.1	Kramerius vyžaduje . . . . .	5
3.2	Struktura datového balíčku . . . . .	6
3.3	Externí odkazy ve FOXML . . . . .	6
3.4	Uložení dat . . . . .	6
3.5	Průběh exportu . . . . .	6
3.6	Omezení procesu importu . . . . .	7
<b>4</b>	<b>Připojované adresáře</b>	<b>9</b>
4.1	Připojované adresáře na Kramerius server . . . . .	9
4.2	Připojované adresáře na Image server . . . . .	9
<b>5</b>	<b>Případy užití</b>	<b>11</b>
5.1	Timeout procesu zpracování FOXML souboru na Kramerius server . . . . .	11
5.2	Import do Kramerius serveru skončil s chybou . . . . .	11
5.3	Zopakování exportu z eDepositů do Krameria . . . . .	12
5.4	Již proběhlo jedno automatické zopakování importu . . . . .	12
<b>6</b>	<b>Příklady kódu</b>	<b>15</b>
6.1	Vytvoření žádosti o export . . . . .	15
6.2	Odeslání žádosti o export . . . . .	15
6.3	Uložení dat k importu do Krameria . . . . .	16
6.4	Smazání importovaných dat v Krameriovi . . . . .	16
<b>7</b>	<b>Vzorky dat</b>	<b>17</b>
7.1	Importní balíček . . . . .	17
7.2	Žádost o export . . . . .	17



---

### Požadavky

---

aplikace:

- přijímá data k importu z aplikace RabbitMQ přes AMQP protokol
- vytváří FOXML soubor
- přijímá náhled první strany ve formátu jpeg2000
- kopíruje importní balíček do Kramerius server přes scp protokol
- odesílá archivní balíček do Storage server pres AMQP protokol
- startuje proces importu voláním REST API na Kramerius server
- zjišťuje průběh importu voláním REST API na Kramerius server
- maže importní balíček na Kramerius server poté, co jej Kramerius server úspěšně naimportuje
- posílá zpátky zprávu o úspěšném importu do aplikace RabbitMQ přes AMQP protokol
- kontroluje, že všechny linky ve FOXML souboru existují
- 1x import zopakuje, jestli dojde k chybě importu



---

### Omezení

---

aplikace:

- použít v `Kramerius server` k importu jen jeden proces najednou. Až jeden import skončí, pustí druhý.





---

## Export do Krameria

---

### 3.1 Kramerius vyžaduje

#### original file

- PDF soubor s ePublikací

#### image source file

- hlavní náhled na ePublikaci
- `preview file` a `thumbnail file` jsou generovány za běhu pomocí `image server`
- soubor je formátu `jpeg2000`

#### preview file

- Kramerius potřebuje alespoň jeden obrázek k zobrazení
- uživatel na něj v Krameriovi klikne
- ... a Kramerius zpřístupní hlavní PDF s ePublikací

---

**Note:** `preview file` je generován na požádání přes `image server`. `FOXML` soubor obsahuje linku na `image server`.

---

#### thumbnail file

- Kramerius soubor používá jako náhled

---

**Note:** soubor je generován stejně jako `preview file` přes `image server`.

---

#### FOXML file

- soubor ve formátu `FOXML`
- obsahuje všechny informace o zobrazovaných souborech
- obsahuje všechny informace o samotné ePublikaci
- obsahuje linky na `original file`, `preview file`, `thumbnail file`
- obsahuje všechny informace o umístění zobrazení ePublikace. Jinak řečeno, obsahuje informace o stromu uložení ePublikace v Krameriovi

---

**Note:** Každý balíček má své `UUID`. Adresář s importními daty je podle něj pojmenovaný. Stejně tak i `FOXML` soubor.

---

## 3.2 Struktura datového balíčku

Archiv obsahuje datové balíčky připravené k importu:

```
jan@jan-XPS-L421X:~/ $ tree e65d9072-2c9b-11e5-99fd-b8763f0a3d61
e65d9072-2c9b-11e5-99fd-b8763f0a3d61
-- edeposit-url.txt
-- e65d9072-2c9b-11e5-99fd-b8763f0a3d61.xml
-- first-page
  -- filename
  -- mimetype
  -- robotandbaby_001.jp2

1 directory, 5 files
```

jednotlivé soubory a adresáře:

e65d9072-2c9b-11e5-99fd-b8763f0a3d61.xml	FOXML soubor
edeposit-url.txt	soubor s odkazem do aplikace eDeposit
first-page	náhled první strany

**Note:** Kramerius dostane jen FOXML soubor. Celý balíček bude uložen v archivu. Přes něj může Kramerius získat další části balíčku.

---

## 3.3 Externí odkazy ve FOXML

Umístění ePublikace v Krameriovi je popsán v sekci `RelSExt`.

Tato sekce musí obsahovat celou cestu k ePublikaci v Krameriovi. Všechny nové větve stromu (cesta k ePublikaci) Kramerius vygeneruje.

## 3.4 Uložení dat

### **edeposit storage**

- poskytuje všechny originály ePublikací
- poskytuje náhled první strany pro `image server`. Ten z něj generuje náhledy.
- poskytuje archiv dat, která byla importována do Krameria

### **kramerius server**

- na tomto serveru jsou uloženy FOXML soubory k importu

## 3.5 Průběh exportu

1. eDeposit Plone aplikace připraví data pro export do Krameria viz *Vytvoření žádosti o export*
2. eDeposit Plone aplikace volá AMQP službu k exportu do Krameria - a poskytne informace k vytvoření FOXML viz *Odeslání žádosti o export*
3. `edeposit.amqp.kramerius` převede MARCXML data do MODS pomocí AMQP služby `edeposit.amqp.marcxml2mods`

4. `edeposit.amqp.kramerius` vytvoří FOXML soubor a celý importní balíček viz *Importní balíček*
5. `edeposit.amqp.kramerius` odešle importní balíček do archivu storage serveru přes AMQP protokol
6. `edeposit.amqp.kramerius` zkopíruje importní balíček na Kramerius server viz *Uložení dat k importu do Krameria*
7. `edeposit.amqp.kramerius` volá REST API na Kramerius server - aplikace spustí na Kramerius server proces importu
8. `edeposit.amqp.kramerius` kontroluje stav importu - aplikace se pravidelně dotazuje na Kramerius server jak proces importu probíhá
9. `edeposit.amqp.kramerius` smaže data k importu - jakmile proces importu úspěšně skončí viz *Smazání importovaných dat v Krameriovi*
10. `edeposit.amqp.kramerius` posílá zprávu s odpovědí - odesílá jí do aplikace RabbitMQ přes AMQP protokol.

## 3.6 Omezení procesu importu

eDeposit může v Krameriovi nastarovat jen jeden proces importu. Poté co proces importu skončí, může nastarovat další.

Data v importním adresáři na serveru Kramerius mohou být uložena tak jak budou přicházet všechna. Čekají na spuštění odpovídajícího importu.



---

## Připojované adresáře

---

Data, která `Kramerius server` vyžaduje, jsou uložena na serveru `Storage server`. Ten je součástí projektu `eDeposit`.

`eDeposit storage server` data poskytuje pomocí `smb` protokolu.

### 4.1 Připojované adresáře na `Kramerius server`

#### `/kramerius_edeposit_import`

- v tomto adresáři jsou uloženy importní balíčky čekající na import do `Krameria`
- `edeposit.amqp.kramerius` je po úspěšném importu smaže

#### `/kramerius_edeposit_storage`

- v tomto adresáři jsou uloženy originály ke zpřístupňování

```
mount -t cifs //10.10.0.42/naki/originals t2 -o username=edeposit,domain=ULTRA_NT  
/kramerius_edeposit_storage
```

#### `/kramerius_edeposit_archive`

- v tomto adresáři jsou uloženy balíčky importovaných dat

```
mount -t cifs //10.10.0.42/naki/archive t2 -o username=edeposit,domain=ULTRA_NT  
/kramerius_edeposit_archive
```

### 4.2 Připojované adresáře na `Image server`

#### `/edeposit_archive`

- v tomto adresáři jsou uloženy importované balíčky včetně náhledu první strany

```
mount -t cifs //10.10.0.42/naki/archive t2 -o username=edeposit,domain=ULTRA_NT  
/edeposit_archive
```



---

## Případy užití

---

### 5.1 Timeout procesu zpracování FOXML souboru na Kramerius server

*účastníci:*

- systém

*vstupní podmínky:*

- objevil se nějaký problém při importu do Krameria
- v importním adresáři jsou odpovídající data k importu

*sled událostí:*

- systém počká 1 den
- pokud problém trvá,
  - systém import zopakuje
- pokud již proběhl 1 pokus o import
  - ePublikace skončí ve stavu “Chyba exportu do Krameria”

*následné podmínky:*

- na serveru Kramerius budou přepsané odpovídající importní data
- na serveru Kramerius bude spuštěn importní proces (pokud není omezení)

### 5.2 Import do Kramerius serveru skončil s chybou

*účastníci:*

- správce Krameria
- správce eDeposit

*vstupní podmínky:*

- objevil se nějaký problém při importu do Krameria
- v importním adresáři jsou odpovídající data k importu

*sled událostí:*

- správce Krameria se obrátí na správce eDeposit
- správce eDeposit si zobrazí detail ePublikace
- správce eDeposit klikne na linku “Zopakovat export do Krameria”
- systém zopakuje export do Krameria

*následné podmínky:*

- na serveru Kramerius budou přepsané odpovídající importní data
- na serveru Kramerius bude spuštěn importní proces (pokud není omezení)

## 5.3 Zopakování exportu z eDeposit do Krameria

*účastníci:*

- správce Krameria
- správce eDeposit

*vstupní podmínky:*

- objevil se nějaký problém při importu do Krameria
- v importním adresáři jsou odpovídající data k importu

*sled událostí:*

- správce Krameria se obrátí na správce eDeposit
- správce eDeposit si zobrazí detail ePublikace
- správce eDeposit klikne na linku “Zopakovat export do Krameria”
- systém zopakuje export do Krameria

*následné podmínky:*

- na serveru Kramerius budou přepsané odpovídající importní data
- na serveru Kramerius bude spuštěn importní proces (pokud není omezení)

## 5.4 Již proběhlo jedno automatické zopakování importu

*účastníci:*

- systém
- správce eDeposit

*vstupní podmínky:*

- objevil se nějaký problém při importu do Krameria
- v importním adresáři jsou odpovídající data k importu
- již proběhl jeden pokus o zopakování importu

*sled událostí:*

- systém přesune ePublikaci do stavu “Chyba exportu do Krameria”
- systém odešle email správci eDeposit



- správce eDepositů se pokusí problém vyřešit

*následné podmínky:*

- na serveru Kramerius budou smazány odpovídající importní data



## Příklady kódu

## 6.1 Vytvoření žádosti o export

```
import json
import base64

marcxml=open('resources/oai_marc.xml','rb').read()
first_page_fname = 'resources/robotandbaby_001.jp2'
location_at_kramerius = "/monografie/2001/John McCarthy/Robot and Baby"
open('resources/export-request.json','wb').write(
    json.dumps(
        dict( __nt_name="ExportToKramerius",
              uuid='e65d9072-2c9b-11e5-99fd-b8763f0a3d61',
              urnnbn='urn:nbn:cz:mzk-0005ol',
              b64_marcxml=base64.encodestring(marcxml),
              first_page = dict(filename=first_page_fname.split('/')[-1],
                                mimetype='image/jp2',
                                b64_data = base64.encodestring(
                                    open(first_page_fname, 'rb').read())),
              location_at_kramerius = location_at_kramerius,
              is_private = False,
            )
    )
)
```

## 6.2 Odeslání žádosti o export

```
import pika
msg = open("resources/export-request.json","rb").read()
conn = pika.BlockingConnection(
    pika.URLParameters("http://guest:guest@localhost:5672/kramerius"))
channel = conn.channel()
channel.basic_publish("export", "request", msg,
                    pika.BasicProperties(content_type="application/json",
                                         delivery_mode=2))
```

## 6.3 Uložení dat k importu do Krameria

```
scp -i ~/.ssh/id_rsa-edeposit-to-kramerius
resources/e65d9072-2c9b-11e5-99fd-b8763f0a3d61/e65d9072-2c9b-11e5-99fd-b8763f0a3d61.xml
edeposit@HAPPKRAM2.nkp.cz:/home/kramerius/kramerius_edeposit_import
```

## 6.4 Smazání importovaných dat v Krameriovi

```
ssh -i ~/.ssh/id_rsa-edeposit-to-kramerius
edeposit@HAPPKRAM2.nkp.cz
"cd /home/kramerius/kramerius_edeposit_import;
rm e65d9072-2c9b-11e5-99fd-b8763f0a3d61.xml"
```

## 7.1 Importní balíček

e65d9072-2c9b-11e5-99fd-b8763f0a3d61.xml

first-page.jp2

edeposit-url.txt

## 7.2 Žádost o export

export-request.json