



RAHVUSARHIIV

Andmebaasi arhiveerimine

Tehniline protsess

Kuldar Aas

Rahvusarhiivi digitaalarhiiv

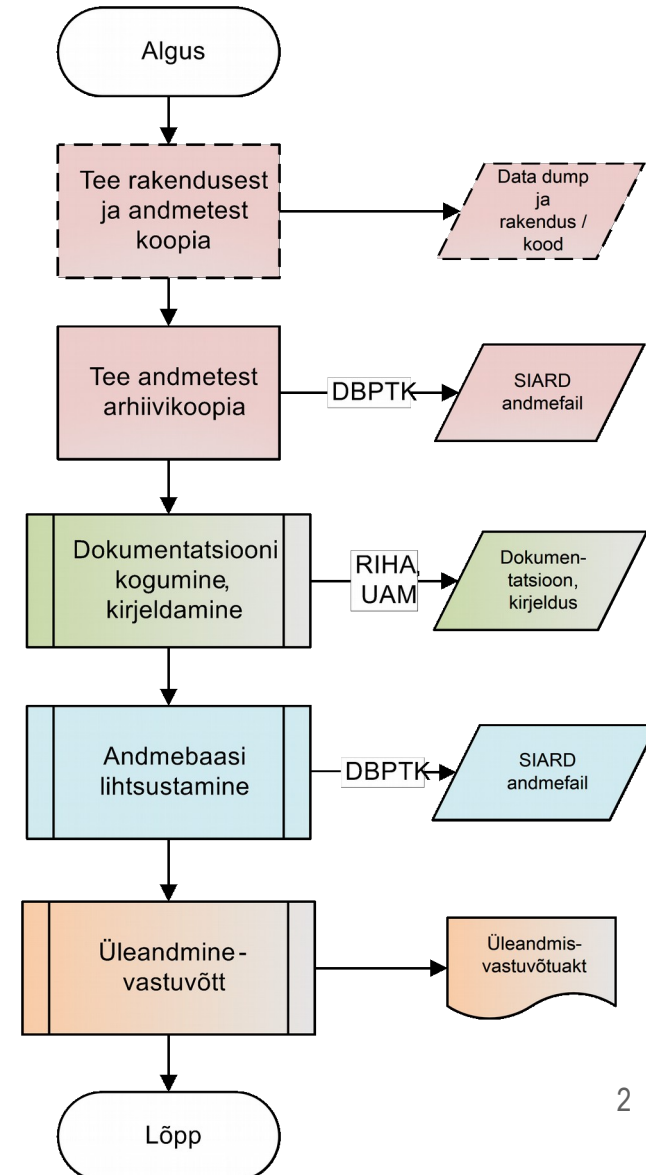
AKde arhiveerimise infopäev

24.01.2018



PROTSESSI SAMMUD

- Arhiveeri andmebaas *as is*, kindla kuupäeva seisuga
 - *Data dump* + rakendus (kui võimalik ja mõistlik)
 - Standardiseeritud SIARD failina (rakendus: DBPTK)
- Dokumentatsiooni ja kirjelduste kogumine
- Lihtsustatud andmebaasi arhiveerimine
 - Standardiseeritud SIARD failina
- Üleandmine / vastuvõtt





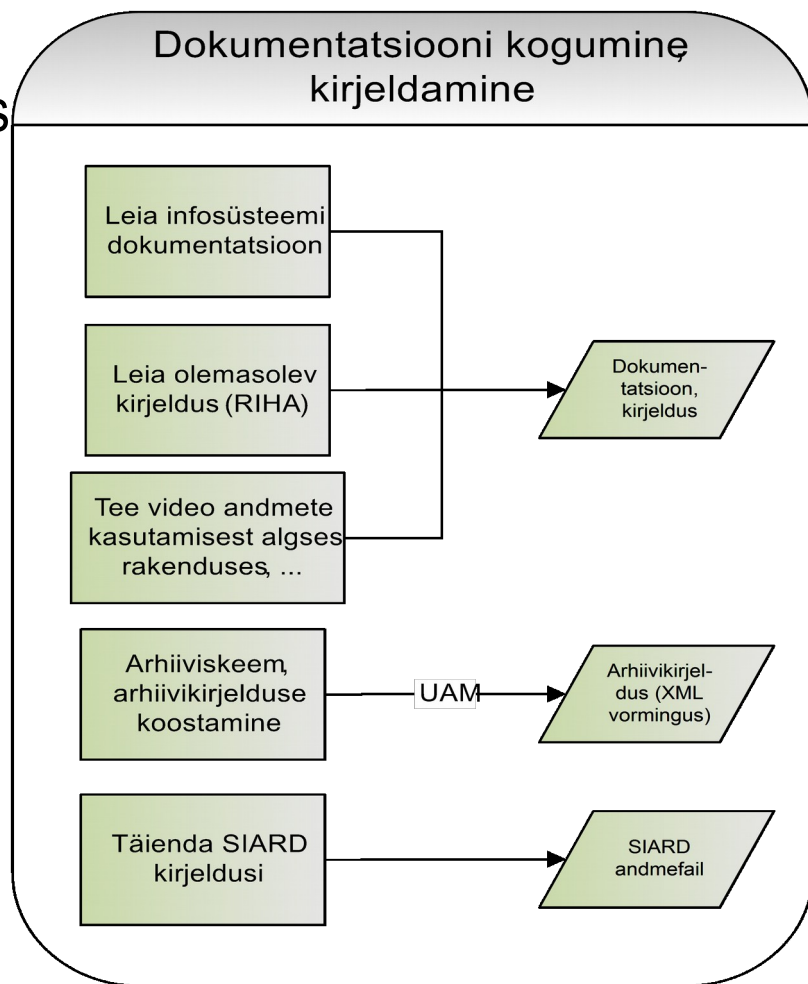
DOKUMENTATSIOON JA KIRJELDUSED

Eesmärk 1: arhiveeritud andmebaasi leitavus arhiivis

- Arhiivikirjelduse loomine
- RIHA kirjelduste arhiveerimine

Eesmärk 2: (arhiivi) kasutaja saab tutvuda andmebaasi sisu, struktuuri ja algse kontekstiga

- Valik olemasolevast dokumentatsioonist (andmemudel, arhitektuur, jne)
- Video süsteemi kasutamisest, jne



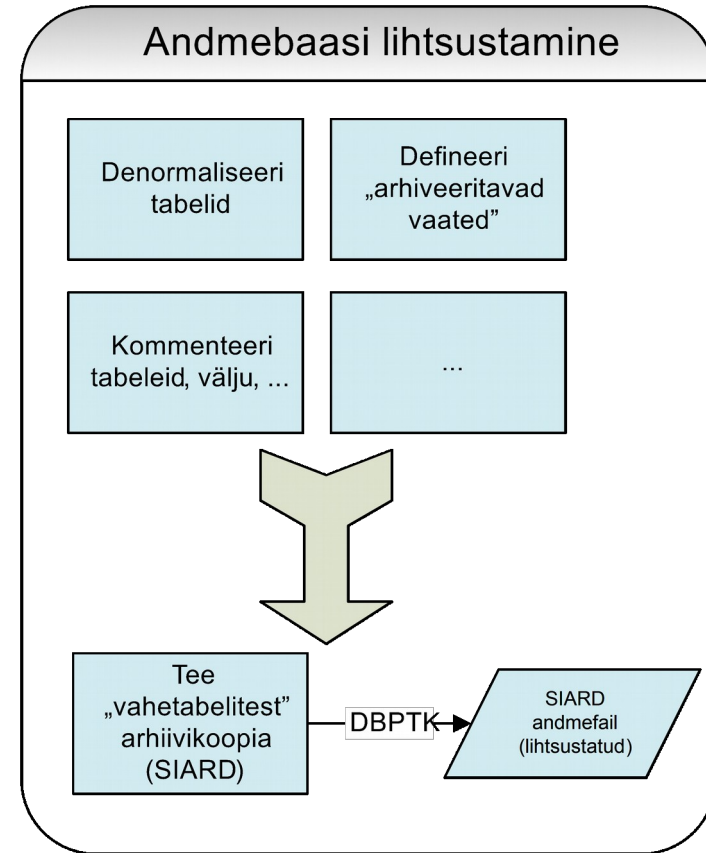


ANDMEBAASI LIHTSUSTAMINE

Eesmärk: arhiveeritud andmebaasi taaskasutamise lihtsustamine

- Tabelite denormaliseerimine
- „Arhiveeritavate vaadete“ defineerimine
- Andmebaasis tabelite, väljade jms kommenteerimine

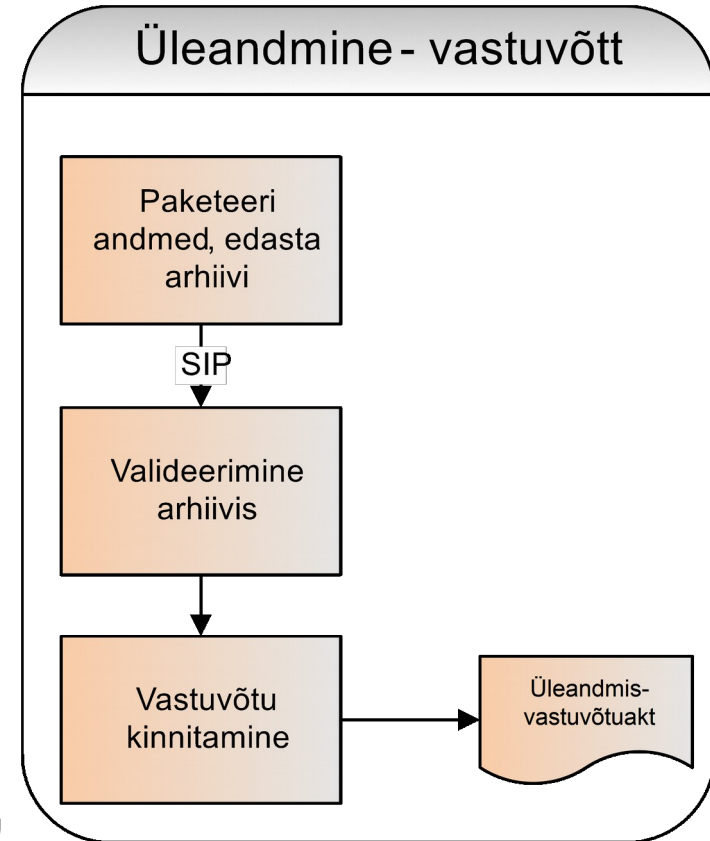
Tulem: Andmebaasi tõmmis lihtsamini kasutataval kujul





ÜLEANDMINE JA VASTUVÕTT

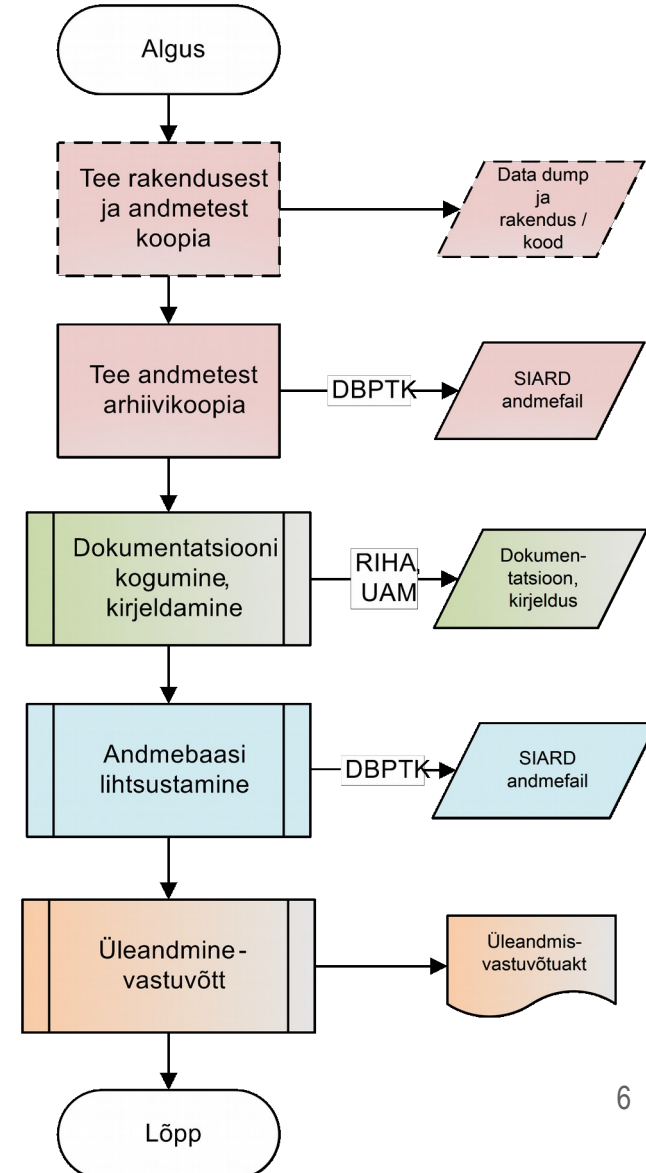
- Kõik loodud komponendid edastatakse arhiivi
 - Data dump, algne SIARD, lihtsustatud SIARD, dokumentatsioon, kirjeldused
- Arhiiv teeb kvaliteedikontrolli
- Arhiiv kinnitab vastuvõtu ametliku aktiga
 - NB! Akti saamise järel on asutusel õigus mittevajalikud andmed hävitada!





- Koostöö!
 - Andmebaasid on erinevad, valikud ja võimalused on erinevad
 - Protsessi detailid lepitakse kokku töö käigus
- „Algne ja autentne“ vs „kasutajale mugav“ vs kulud
- Vormingud ja tööriistad: SIARD, DBPTK, (DBVTK)

KOKKUVÕTE





ANDMEBAASI ARENDAMINE

Soovitused lihtsamaks arhiveerimiseks

- Andmebaasi kommentaarid (baasi sees)
 - Tabelid, väljad / veerud, vaated, protseduurid / funktsioonid
 - Selge ja täpne olemi kirjeldus
 - Juurdepääsupiirangute olemasolu
 - Kui andmebaasi uuendatakse, siis uuenda ka kommentaar!
- Andmebaasipäringute jaoks kasuta protseduure või view'sid (vt ka ISKE M2.134)
- Võimalusel erista andmemudelid juurdepääsupiiranguga andmed avalikest
- Võimalusel erista andmemudelid tehnilised, nn „tugitabelid“ päris andmetest
- Nimekiri hakkab aja jooksul täienema ..



SIARD

- Software Independent Archiving of Relational Databases
- SIARD 1.0 loodud Šveitsi Liiduarhiivi poolt 2003 – 2004
- Avatud vorming andmebaaside sisu jaoks
 - aluseks XML, SQL:99, UTF-8
- Šveitsi föderaaltasandi standard andmebaaside arhiveerimiseks (eCH-0165)
- Laialt kasutusel ka teistes riikides
- 2014-2015 uuendamine E-ARKi raames → SIARD 2.0
 - Parema skaleeruvuse, SQL:2008 toetus



The screenshot shows a Windows Explorer window titled "C:\Documents and Settings\u1984\Desktop\tmp". The address bar shows the path "C:\Documents and Settings\u1984\Desktop\tmp". The main pane displays a table with the following data:

Name	Größe	Typ	Geändert am
LADIS.siard	28,286 KB	SIARD-Datei	9/2/2009 2:43 PM

A second Explorer window is open over the "LADIS.siard" file, showing its contents. The address bar of this window shows ".Desktop\tmp". The main pane displays a table with the following data:

Name	Größe	Typ	Geändert am
content		File Folder	9/7/2009 11:32 AM
header		File Folder	9/7/2009 11:32 AM

The status bar at the bottom of the Explorer window shows "1 Objekt(e)", "27.6 MB", and "My Computer".



The screenshot shows a Windows Explorer window titled "C:\Documents and Settings\u1984\Desktop\tmp". The address bar shows the path "C:\Documents and Settings\u1984\Desktop\tmp". The main pane displays a table with the following data:

Name	Größe	Typ	Geändert am
LADIS.siard	28,286 KB	SIARD-Datei	9/2/2009 2:43 PM

On the left sidebar, under "Andere Orte", the "Desktop" location is selected. Below it, a smaller Explorer window titled "Desktop\tmp" is open, showing a table with the following data:

Name	Größe	Typ	Geändert am
content		File Folder	9/7/2009 11:32 AM
header		File Folder	9/7/2009 11:32 AM

The "Details" pane on the left shows information for the "tmp" folder: "File Folder", "Geändert am: Monday, September 07, 2009, 2:19 PM". Below the details, a context menu is open, listing "metadata.xml", "metadata.xsd", and "metadata.xsl".

At the bottom of the Explorer window, the status bar shows "1 Objekt(e)", "27.6 MB", and "My Computer".



The screenshot shows a Windows Explorer window titled "C:\Documents and Settings\u1984\Desktop\tmp". The address bar shows the path "C:\Documents and Settings\u1984\Desktop\tmp". The main pane displays a table of files:

Name	Größe	Typ	Geändert am
LADIS.siard	28,286 KB	SIARD-Datei	9/2/2009 2:43 PM

Below this, a smaller window shows the contents of the "tmp" folder:

Name	Größe	Typ	Geändert am
content		File Folder	9/7/2009 11:32 AM
header		older	9/7/2009 11:32 AM
schema0		older	9/7/2009 11:32 AM
table0		older	
table1		older	
table2		older	
table3		older	
table4		older	
table5		older	
table6		older	
table7		older	

Arrows indicate the following actions:

- An arrow points from the "LADIS.siard" file in the top window to the "tmp" folder in the bottom window.
- An arrow points from the "header" folder in the bottom window to a context menu.
- Another arrow points from the "content" folder in the bottom window to the same context menu.

The context menu shows the following items:

- metadata.xml
- metadata.xsd
- metadata.xsl

The status bar at the bottom indicates "1 Objekt(e)", "27.6 MB", and "My Computer".



The screenshot shows a Windows Explorer window titled "C:\Documents and Settings\u1984\Desktop\tmp". The address bar shows the path "C:\Documents and Settings\u1984\Desktop\tmp". The main pane displays a table of files:

Name	Größe	Typ	Geändert am
LADIS.siard	28,286 KB	SIARD-Datei	9/2/2009 2:43 PM

Below this, a smaller window shows the contents of the "tmp" folder:

Name	Größe	Typ	Geändert am
content		File Folder	9/7/2009 11:32 AM
header		older	9/7/2009 11:32 AM
schema0		older	9/7/2009 11:32 AM

The "content" folder is expanded, showing a list of subfolders:

- table0
- table1
- table2
- table3
- table4
- table5
- table6
- table7

The "table3" folder is expanded, showing files:

- lob5
- table37.xml
- table37.xsd

Arrows indicate the navigation path: from the "LADIS.siard" file in the top window to the "tmp" folder in the middle window, then to the "content" folder, then to the "table3" folder, and finally to the "table37.xsd" file.

The "Details" pane on the left shows information for the "tmp" folder:

tmp
File Folder
Geändert am: Monday, September 07, 2009, 2:19 PM

Below the details, a preview of files is shown:

- metadata.xml
- metadata.xsd
- metadata.xsl

The status bar at the bottom shows "1 Objekt(e)", "27.6 MB", and "My Computer".



- Database Preservation Toolkit (DBPTK)
 - Vabavaraline lahendus andmebaasi sisu salvestamiseks SIARD vormingusse
- Database Visualisation Toolkit (DBVTK)
 - Vabavaraline lahendus SIARD vormingus salvestatud andmebaasi vaatamiseks
- www.database-preservation.com



RAHVUSARHIIV

TÄNAN TÄHELEPANU EEST!