Welcome to this live webinar on Archiving databases: Why, what and how?

Start 10:00

20 June 2025

Audience notes for the Live Webinar

Your **cameras have been turned off** and **microphones muted**.

If you have any technical issues during the event, please use the chat function.

Please **use the Q&A for questions to speakers**. These will be addressed at the end of the event.

Please note that this webinar is recorded. No attendee personal information will be captured in these recordings.

Agenda

10:00 – 10:05
eArchiving Initiative welcome
Jaime Kaminski – eArchiving Initiative

10:05 – 10:50
Archiving databases: Why, what and how?
Koit Saarevet, National Archives of Estonia

10:50 – 11:00
Q&A

Archiving databases: Why, what and how?

Koit Saarevet, National Archives of Estonia

eArchiving Initiative Training Webinar

Contents

1. Legal
2. Database models
3. What to archive?
4. Process
5. Descriptions
6. Tools

European Commission

LEGAL

Archives Act

- Subsection 2(1): A record is information recorded on any medium /…/.
- Subsection 2(2): Archival records are records to which archival value has been granted by a public archives as the result of an appraisal.
- Subsection 2(4): Archival records created or received in the course of performance of public duties shall be transferred to the National Archives /…/.

https://www.riigiteataja.ee/en/eli/521032019019/consolide

European Commission

Archives Act

- Subsection 2(1): A record is information recorded on any medium /…/.
- Subsection 2(2): Archival records are records to which **archival value has been granted by a public archives as the result of an appraisal.**
- Subsection 2(4): Archival records created or received in the course of performance of public duties shall be transferred to the National Archives /…/.

https://www.riigiteataja.ee/en/eli/521032019019/consolide

European Commission

NAE appraisal decisions

- HO 51/2017: Identification of databases with archival value
- Institutional appraisals, where AV is given to a specific database
- Other appraisals, if a series that received AV in them has moved to a database

NAE appraisal decisions:

https://www.ra.ee/arhiivihaldus/hindamisotsused/

78

https://www.ra.ee/astra/site/databases

DATABASE MODELS

Database models

- Relational
- Tree
- Network
- Object
- Document
- Star
- …

} NoSQL

European Commission

SQL = Structured Query Language

```
SELECT * FROM city;
```

NoSQL

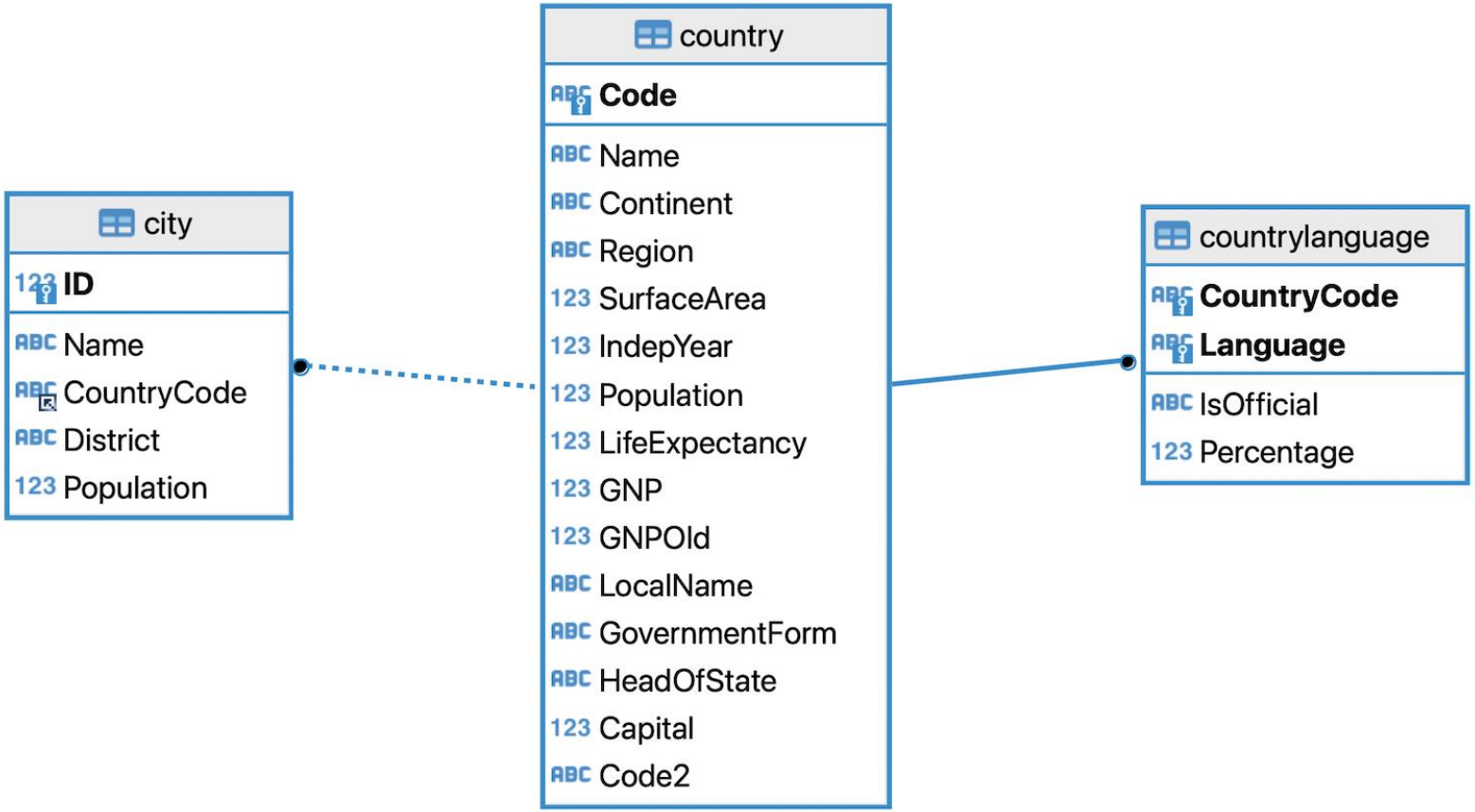
Non-SQL → non-relational

Not only SQL

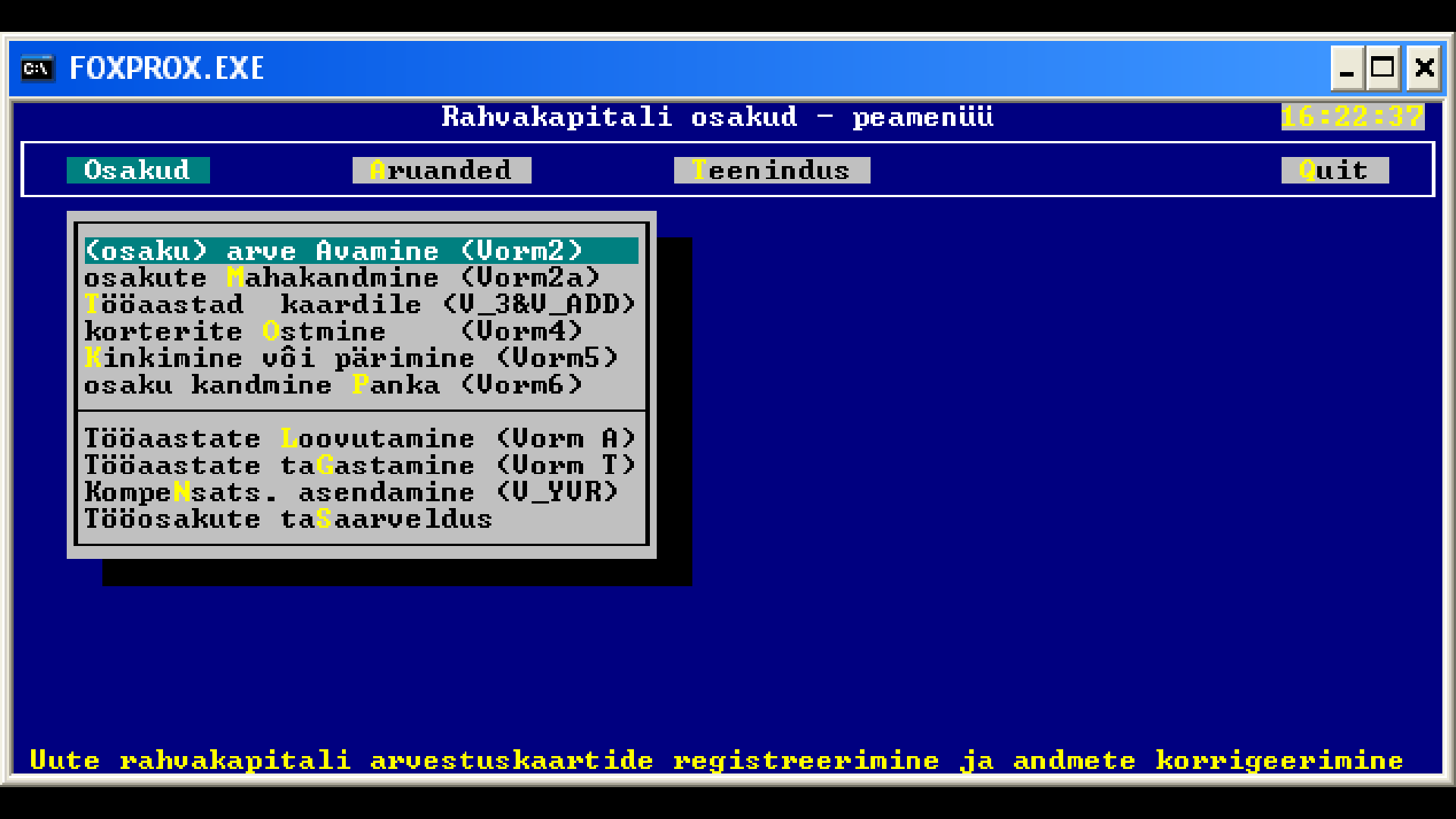
Relational database

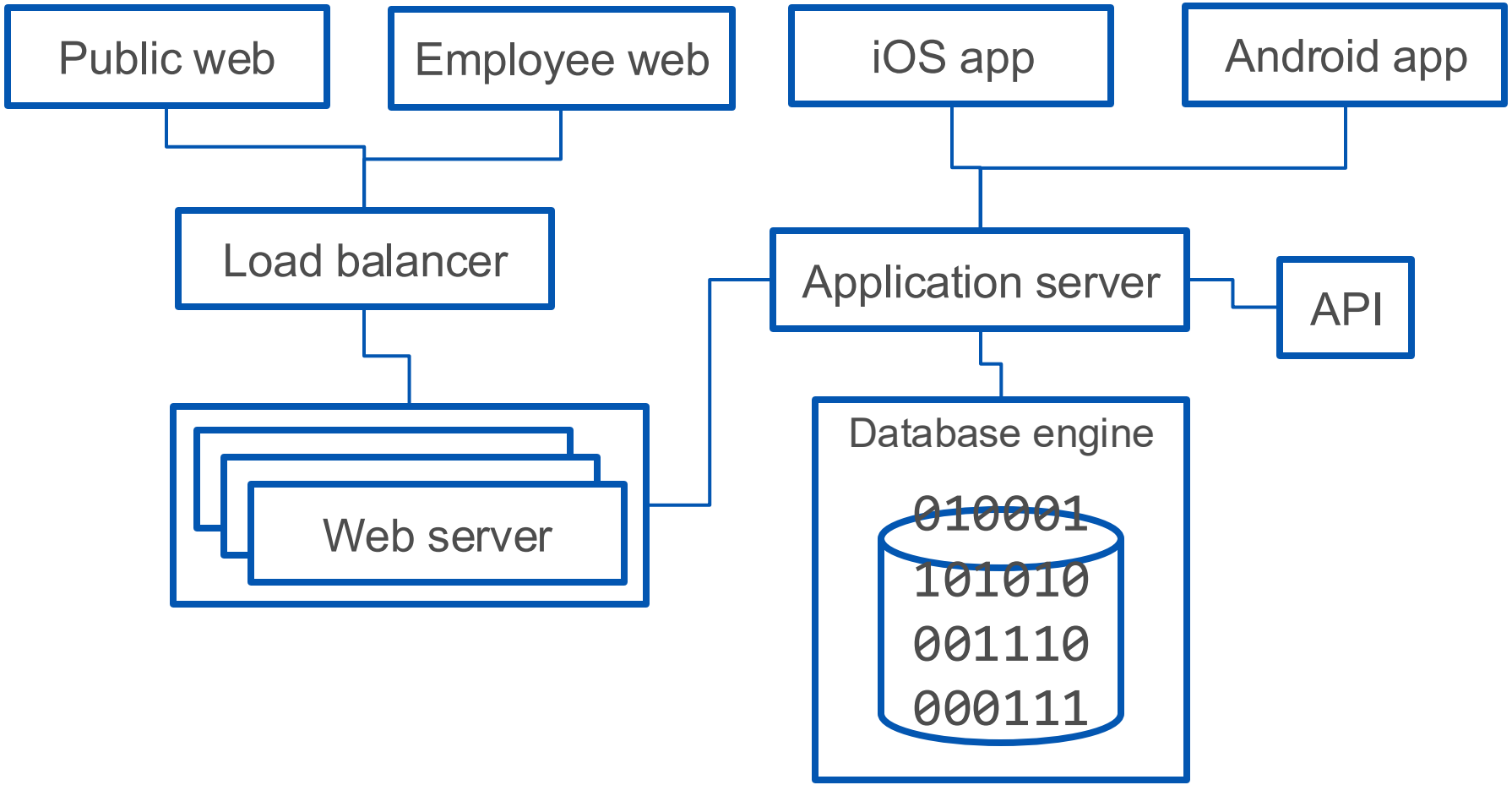
- Data is stored in tables and columns
- Relationships between tables are through the values of data fields
- … i.e. if two tables are related, they must have a column with the same meaning

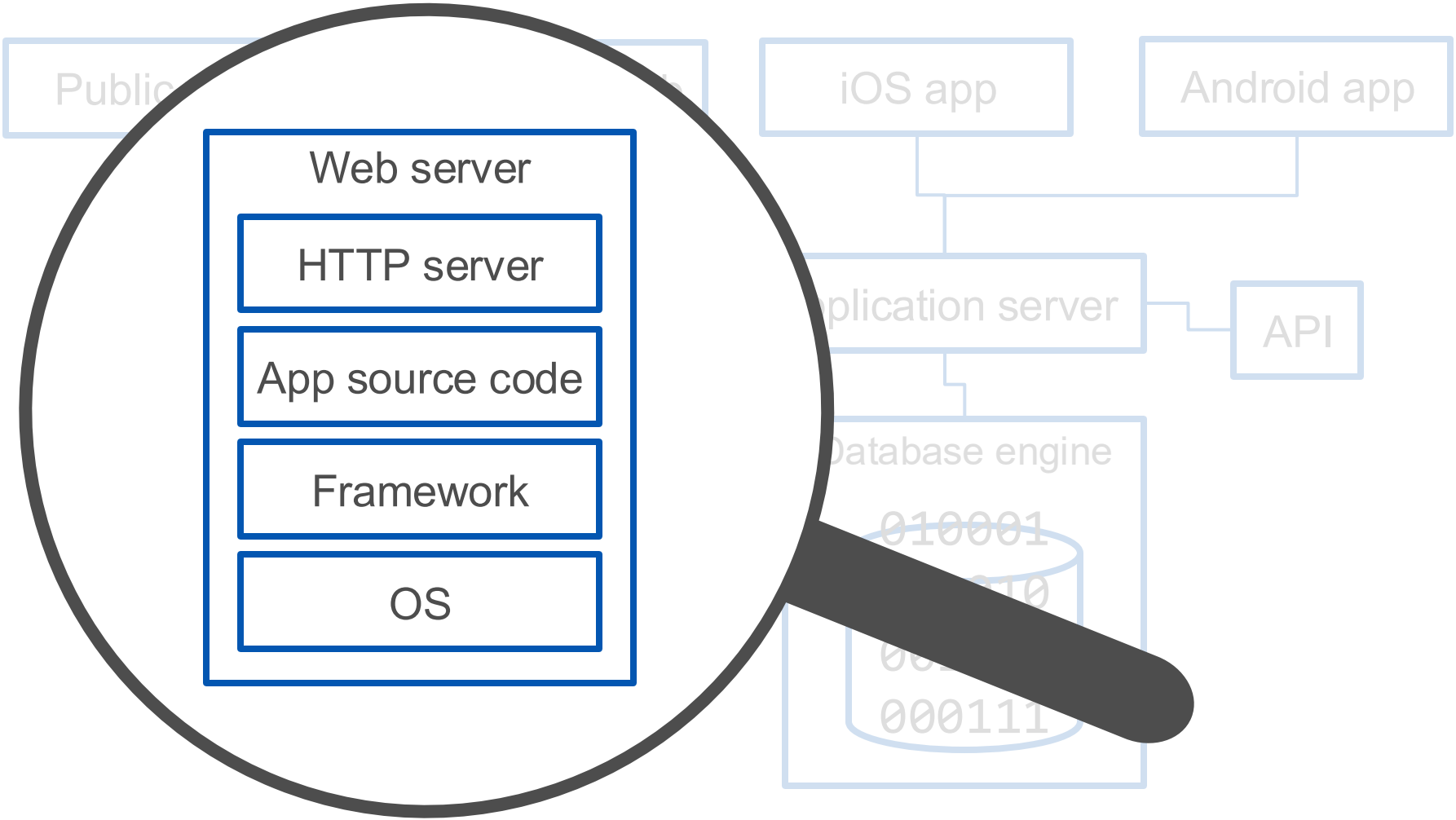
European Commission

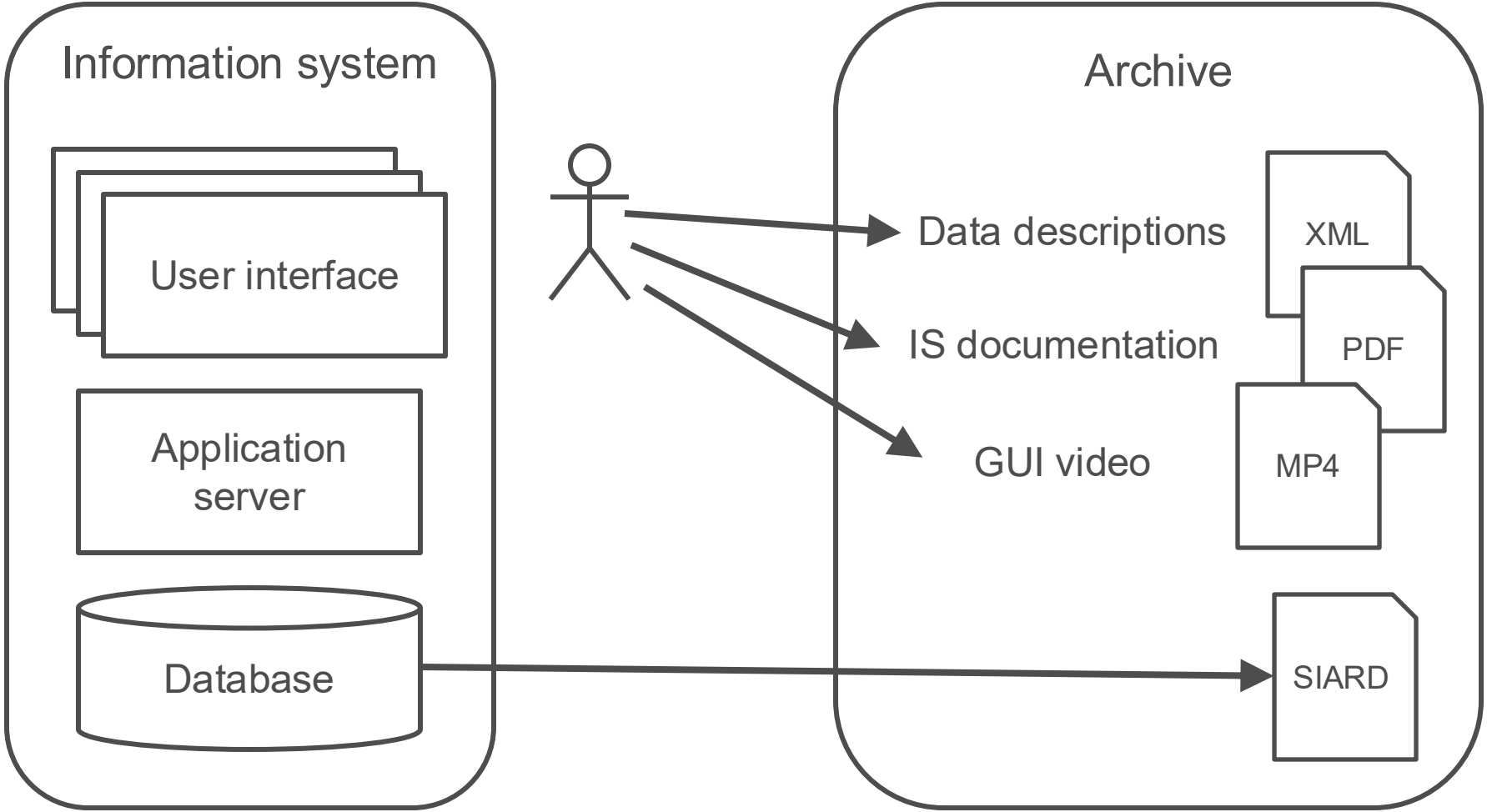

city
ID
Name
CountryCode
District
Population
country
Code
Name
Continent
Region
SurfaceArea
IndepYear
Population
LifeExpectancy
GNP
GNPOld
LocalName
GovernmentForm
HeadOfState
Capital
Code2
countrylanguage
CountryCode
Language
IsOfficial
Percentage

WHAT TO ARCHIVE?


FOXPROX.EXE
Rahvakapitali osakud - peamenüü
16:22:37
Osakud
Aruanded
Teenindus
Quit
(osaku) arve Avamine (Vorm2)
osakute Mahakandmine (Vorm2a)
Tööaastad kaardile (V_3&V_ADD)
korterite Ostmine (Vorm4)
Kinkimine või pärimine (Vorm5)
osaku kandmine Panka (Vorm6)
Tööaastate Loovutamine (Vorm A)
Tööaastate taGastamine (Vorm T)
KompeNsats. asendamine (V_YVR)
Tööosakute taSaarveldus
Uute rahvakapitali arvestuskaartide registreerimine ja andmete korrigeerimine

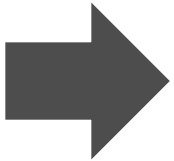

Public web
Employee web
iOS app
Android app
Load balancer
Application server
API
Web server
Database engine
010001
101010
001110
000111


Public
iOS app
Android app
plication server
API
Database engine
010001
000111
Web server
HTTP server
App source code
Framework
OS


Information system
User interface
Application server
Database
Archive
Data descriptions
IS documentation
GUI video
XML
PDF
MP4
SIARD

SIARD =
Software Independent Archiving of Relational Databases

010001
101010
001110
000111



```
<siardArchive>
    <dbname>FR</dbname>
    <description>Forest registry
    </description>
    <archiver>Koit Saarevet
    </archiver>
    …
</siardArchive>
```

SIARD

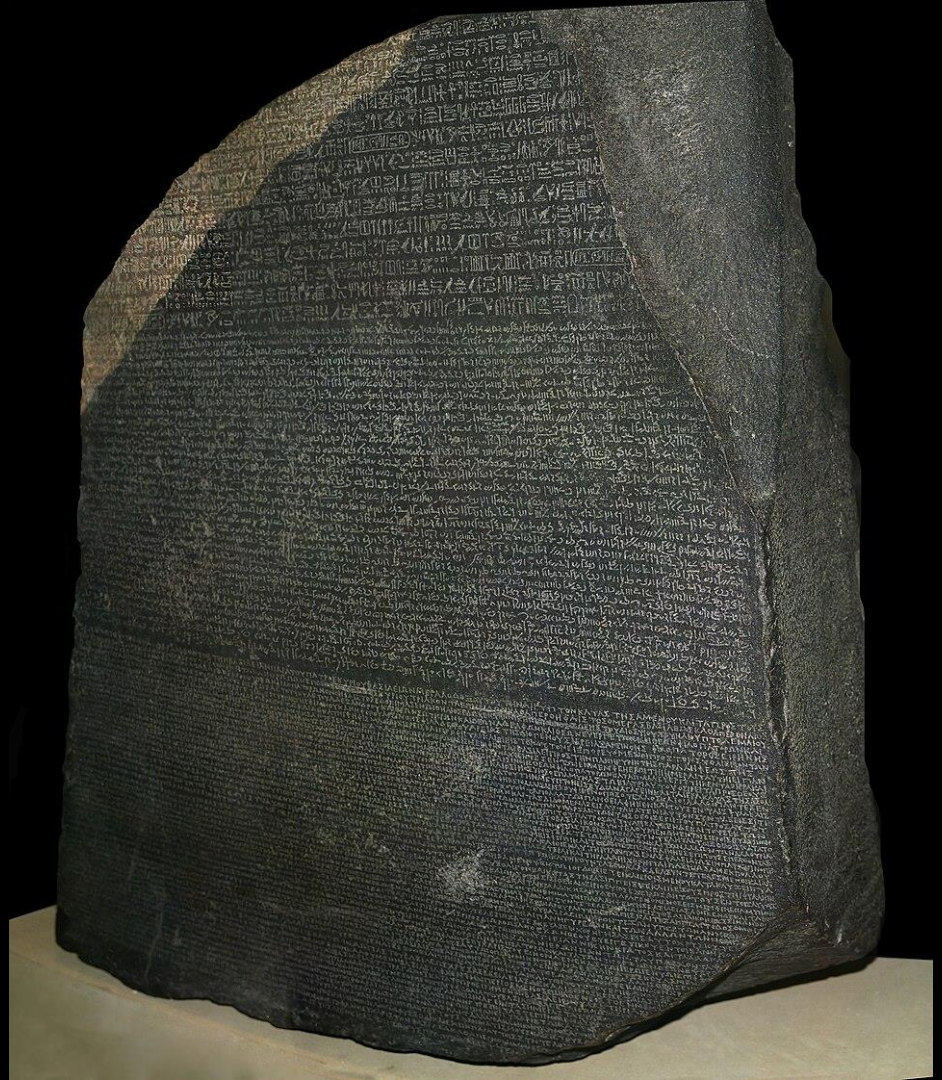
- De facto standard for archiving relational databases (used in most government archives worldwide)
- First issued in 2004, major updates in 2008 and 2016, smaller ones 2019 and 2021
- https://dilcis.eu/content-types/siard

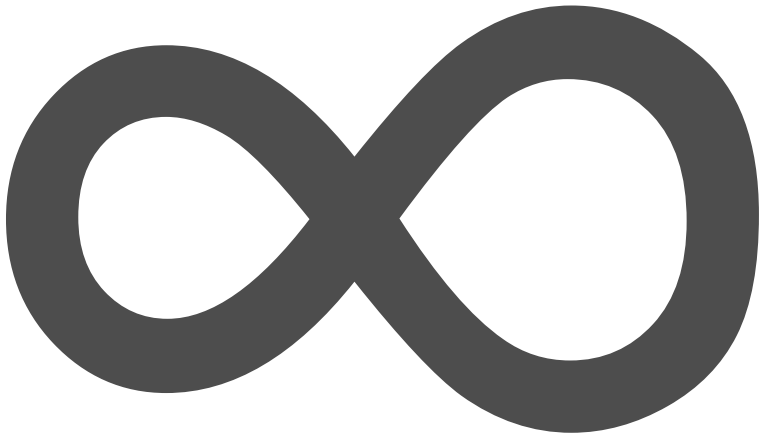
European Commission

PROCESS

1. Initiation
2. Kick-off
3. Cloning the DB
4. Test snapshot 1
5. Appraisal
6. Test snapshot 2
7. Final snapshot
8. Documentation
9. Finalising

DESCRIPTIONS





Archival descriptions

- Title
- Dates
- Archival history
- Scope and content
- …

European Commission

Technical documentation

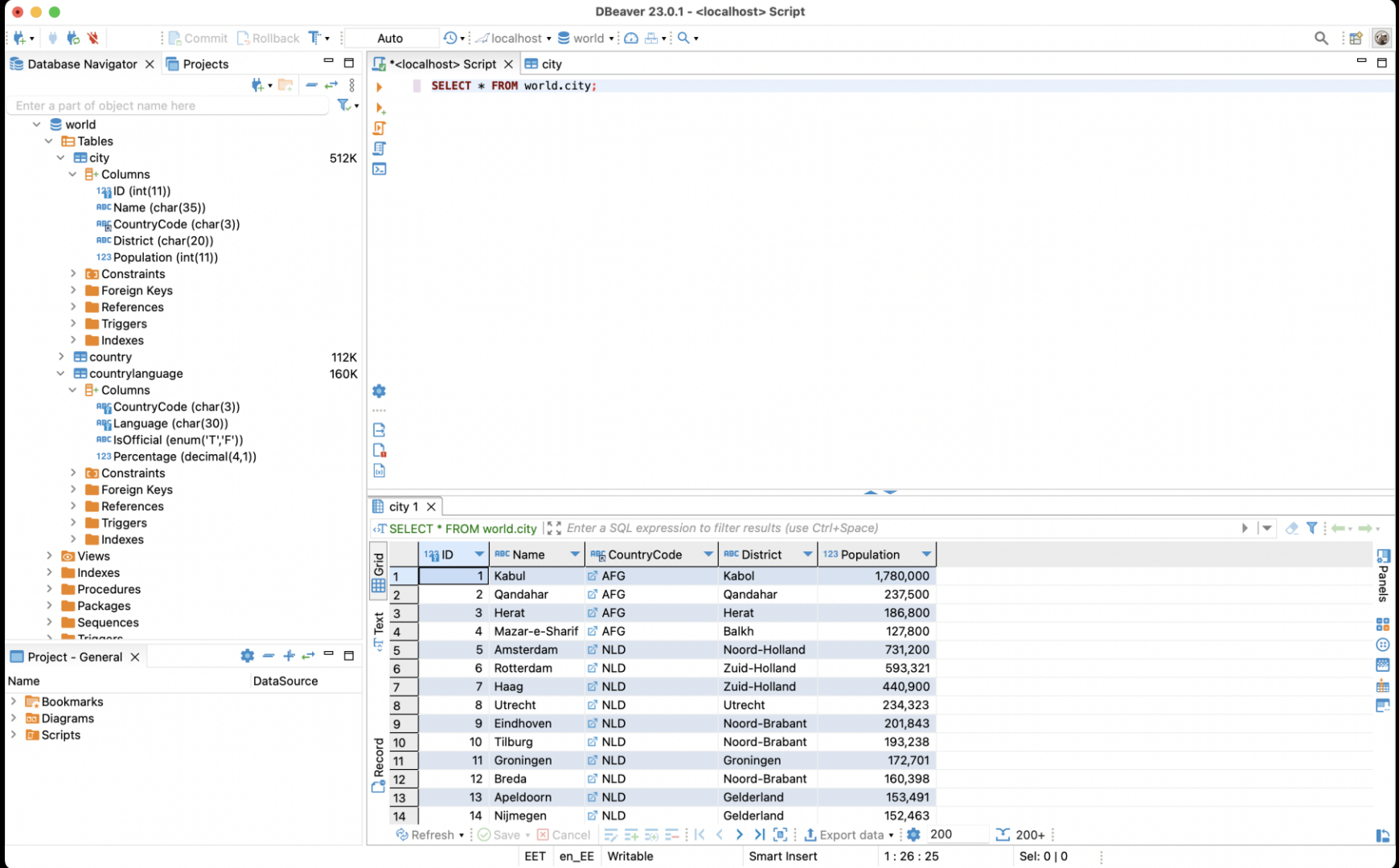
- Data model (conceptual, logical, physical)
  - Table and column descriptions
  - Entity-Relationship Diagram
- Data descriptions (semantics, ontologies)
- Architecture
- User guides
- …

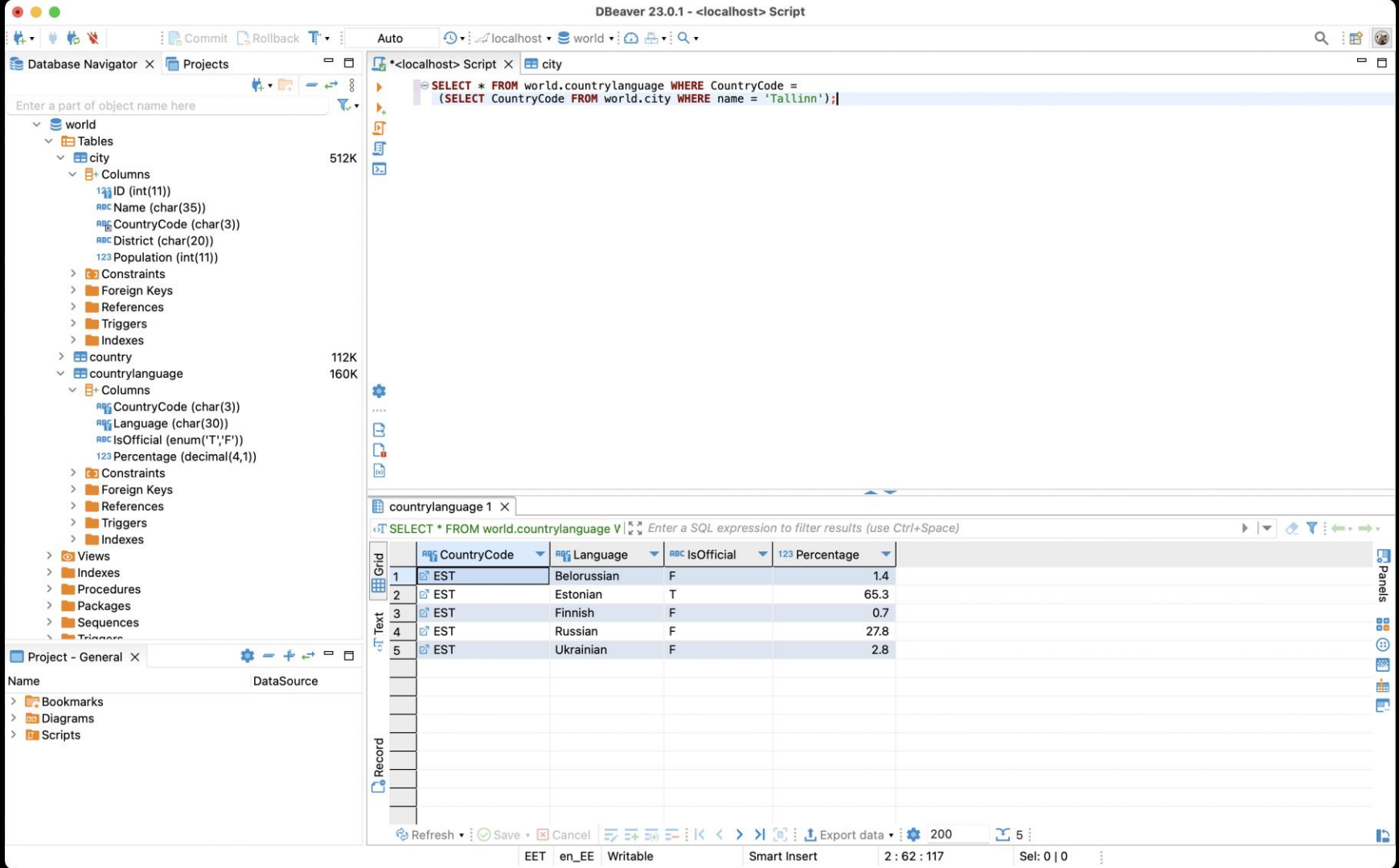
European Commission

6

TOOLS

DBeaver


DBeaver 23.0.1 - <localhost> Script
Commit
Rollback
Auto
localhost
world
Database Navigator
Projects
Enter a part of object name here
world
Tables
city 512K
Columns
ID (int(11))
Name (char(35))
CountryCode (char(3))
District (char(20))
Population (int(11))
Constraints
Foreign Keys
References
Triggers
Indexes
country 112K
countrylanguage 160K
Columns
CountryCode (char(3))
Language (char(30))
IsOfficial (enum('T','F'))
Percentage (decimal(4,1))
Constraints
Foreign Keys
References
Triggers
Indexes
Views
Indexes
Procedures
Packages
Sequences
Project - General
Name
DataSource
Bookmarks
Diagrams
Scripts
*<localhost> Script
city
SELECT * FROM world.city;
city 1
SELECT * FROM world.city
Enter a SQL expression to filter results (use Ctrl+Space)
Grid
Text
Record
Panels
ID
Name
CountryCode
District
Population
1 1 Kabul AFG Kabol 1,780,000
2 2 Qandahar AFG Qandahar 237,500
3 3 Herat AFG Herat 186,800
4 4 Mazar-e-Sharif AFG Balkh 127,800
5 5 Amsterdam NLD Noord-Holland 731,200
6 6 Rotterdam NLD Zuid-Holland 593,321
7 7 Haag NLD Zuid-Holland 440,900
8 8 Utrecht NLD Utrecht 234,323
9 9 Eindhoven NLD Noord-Brabant 201,843
10 10 Tilburg NLD Noord-Brabant 193,238
11 11 Groningen NLD Groningen 172,701
12 12 Breda NLD Noord-Brabant 160,398
13 13 Apeldoorn NLD Gelderland 153,491
14 14 Nijmegen NLD Gelderland 152,463
Refresh
Save
Cancel
Export data
200
200+
EET
en_EE
Writable
Smart Insert
1 : 26 : 25
Sel: 0 | 0

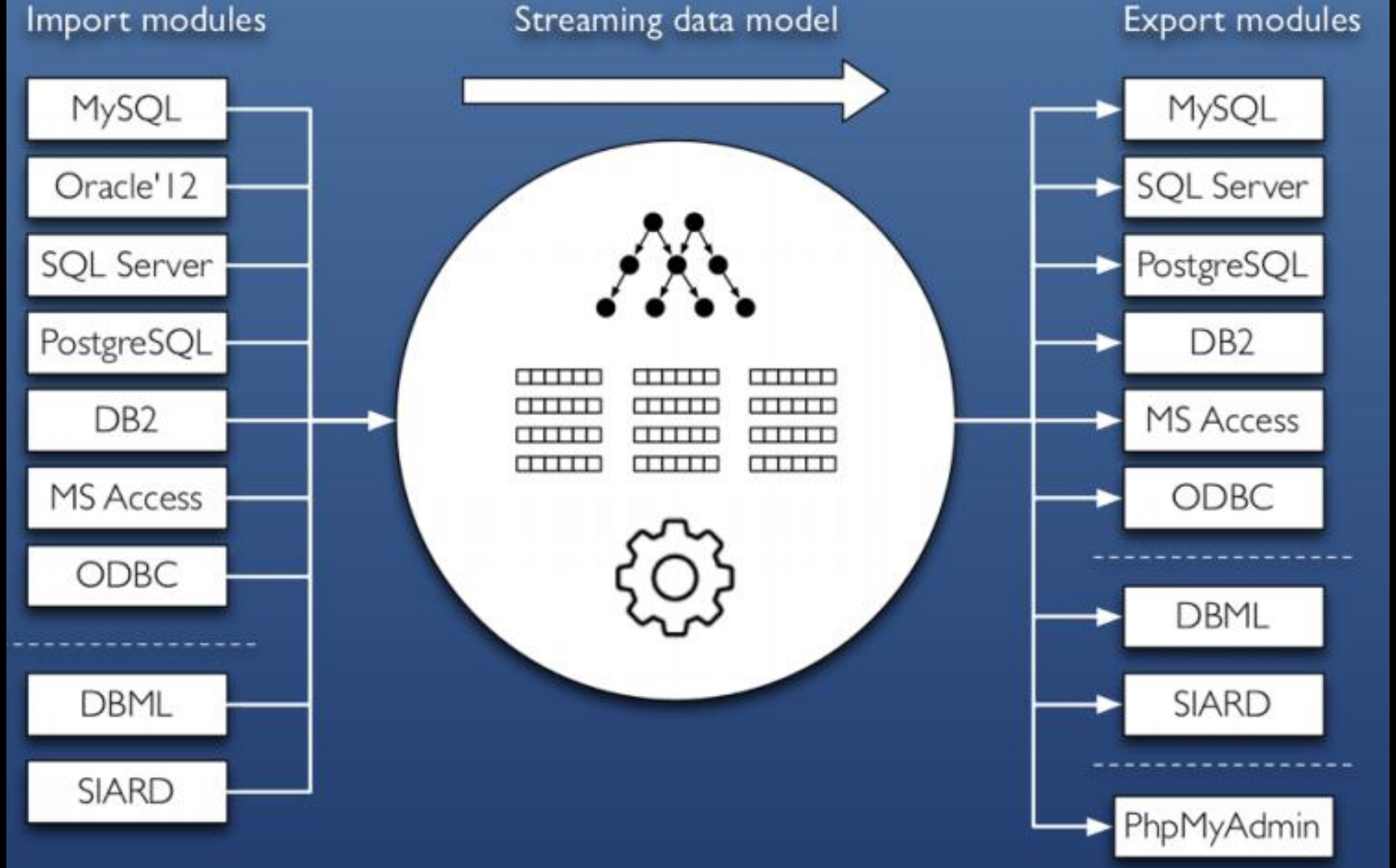

DBeaver 23.0.1 - <localhost> Script
Commit
Rollback
Auto
localhost
world
Database Navigator
Projects
Enter a part of object name here
world
Tables
city 512K
Columns
ID (int(11))
Name (char(35))
CountryCode (char(3))
District (char(20))
Population (int(11))
Constraints
Foreign Keys
References
Triggers
Indexes
country 112K
countrylanguage 160K
Columns
CountryCode (char(3))
Language (char(30))
IsOfficial (enum('T','F'))
Percentage (decimal(4,1))
Constraints
Foreign Keys
References
Triggers
Indexes
Views
Indexes
Procedures
Packages
Sequences
Project - General
Name
DataSource
Bookmarks
Diagrams
Scripts
*<localhost> Script
city
SELECT * FROM world.countrylanguage WHERE CountryCode =
(SELECT CountryCode FROM world.city WHERE name = 'Tallinn');
countrylanguage 1
SELECT * FROM world.countrylanguage V
Enter a SQL expression to filter results (use Ctrl+Space)
Grid
Text
Record
Panels
CountryCode
Language
IsOfficial
Percentage
1 EST Belorussian F 1.4
2 EST Estonian T 65.3
3 EST Finnish F 0.7
4 EST Russian F 27.8
5 EST Ukrainian F 2.8
Refresh
Save
Cancel
Export data
200
5
EET
en_EE
Writable
Smart Insert
2 : 62 : 117
Sel: 0 | 0

https://database-preservation.com

Database Preservation Toolkit (DBPTK)

- Open-source solution for
  - Creating SIARD files
  - Simple full-text search of data stored in a SIARD file
  - Visualisation of database structure
- Developed by KEEP Solutions (Portugal)
  - https://www.database-preservation.com/
  - https://www.keep.pt/en

European Commission


Import modules
Streaming data model
Export modules
MySQL
Oracle'12
SQL Server
PostgreSQL
DB2
MS Access
ODBC
DBML
SIARD
MySQL
SQL Server
PostgreSQL
DB2
MS Access
ODBC
DBML
SIARD
PhpMyAdmin

Create a SIARD file of „world“ database

- Open dbptk-desktop
- Select „Create“
- Select „MySQL“
  - Hostname: localhost
  - Port number: /do not change/
  - Username: /enter your database admin user name/
  - Password: /enter your database admin password/
  - Database: world
  - Disable Encryption: yes
- Follow the wizard, create the database snapshot!

European Commission

DBPTK Developer command line

- Instructions at: https://github.com/keeps/dbptk-developer/wiki/Application-usage

European Commission

Thank you

Contact

https://www.e-ark-foundation.eu

support@e-ark-foundation.eu

https://www.linkedin.com/groups/8343650

https://bsky.app/profile/eu-earchiving.bsky.social

https://www.youtube.com/@e-ark