Practical Database Preservation

Download resources

Wifi: database preservation workshop

Pass: database

http://provided_address

http://192.168.2.1

Luis Faria lfaria@keep.pt
Bruno Ferreira bferreira@keep.pt

DLM 8th Triennial Conference, September 15th 2017, Brighton, UK

Why do we archive databases?

Databases hold a lot of important information

They support organisations' internal information systems

A record of the interactions between the public sector and tax payers

May be used to store scientific data

They hold the records that do not exist in any other form

They keep record day to day business activities

Legal or patrimonial reasons

Legislation in some countries mandates that databases are periodically archived

Financial reasons

Maintenance costs, license fees

Efficient way to preserve a large number of records

Little manual effort is involved when compared to archiving record by record

BAPTISADOS

N.º 10

Maria Ilena

N.º 1 - [illegible] viuva por lhe ter fallecido seu marido José Monteiro Sousa no dia 17 de Dezembro do ano corrente. Registo de obito n.º 394 desta Conservatoria. Baião, 17 de Dezembro de 1956. O Conservador

N.º 2 - Casou com José Monteiro Sousa nesta Repartição do Registo Civil de Baião em 8 de Dezembro de 1921. Assento n.º 70 de 1921. Em 23 de Outubro de 1972. O Conservador José [illegible]

N.º 3 - Faleceu em 6 de Fevereiro de 1950, na freguesia de Gove, deste concelho. Assento n.º 37. Ano 1950. Em 6 de Fevereiro de 1950.

Aos vinte e cinco dias do mez de Fevereiro do anno de mil novecentos n'esta Egreja parochial de São Bartholomeu de Campello, concelho de Baião, diocese do Porto, eu o presbytero Joaquim Antonio Pinto Teixeira, parocho da mesma freguezia baptisei solemnemente um individuo do sexo feminino a quem dei o nome de Maria Ilena e que nasceu n'esta freguezia lugar de Jugilde ás nove horas da noite do dia dezeseis do mez de Fevereiro do anno de mil novecentos filha legitima de Arthur de Azevedo Pinto e Anna de Jesus, proprietarios, naturaes, recebidos, parochianos e moradores no lugar do Jugilde d'esta referida freguezia

neta paterna de João de Azevedo Pinto e Josefa Candida

materna de Manoel Pinto Teixeira e Joaquina de Jesus

Foi padrinho Francisco de Souza Carneiro e madrinha Joaquina de Jesus, madrinha Balbina de Jesus proprietarios do lugar de Jugilde d'esta mencionada freguezia

os quaes todos sei serem os proprios. E para constar se lavrou em duplicado este assento, que, depois de ser lido e conferido perante os padrinhos, assigno só porque os padrinhos disseram que não sabiam escrever. Era ut supra

O parocho Joaquim Antonio Pinto Teixeira

Porto—Typ. a vapor da Officina de S. José.

6 Carvalhos

BAPTISADOS

N.º 11

Joaquim

Falleceu a 24 de Agosto de 1903

Confere o obito

Aos quatro dias do mez de Março do anno de mil novecentos n'esta Egreja parochial de São Bartholomeu de Campello, concelho de Baião, diocese do Porto, eu o presbytero Joaquim Antonio Pinto Teixeira, parocho da mesma freguezia baptisei solemnemente um individuo do sexo masculino a quem dei o nome de Joaquim e que nasceu n'esta freguezia lugar de Amarelhe ás sete horas da manhã do dia vinte e cinco do mez de Fevereiro do anno de mil novecentos filho legitimo de Manoel Pinto Carneiro e Maria da Graça proprietarios do lugar de Amarelhe, naturaes, recebidos e parochianos d'esta dita freguezia.

neto paterno de José Pinto Carneiro e Delfina Rosa de Jesus

materno de Bernardo Teixeira de Queiroz e Thereza de Jesus

Foi padrinho Joaquim Nicolau de Miranda e madrinha Emilia de Jesus proprietarios do lugar de Villares d'esta mesma freguezia

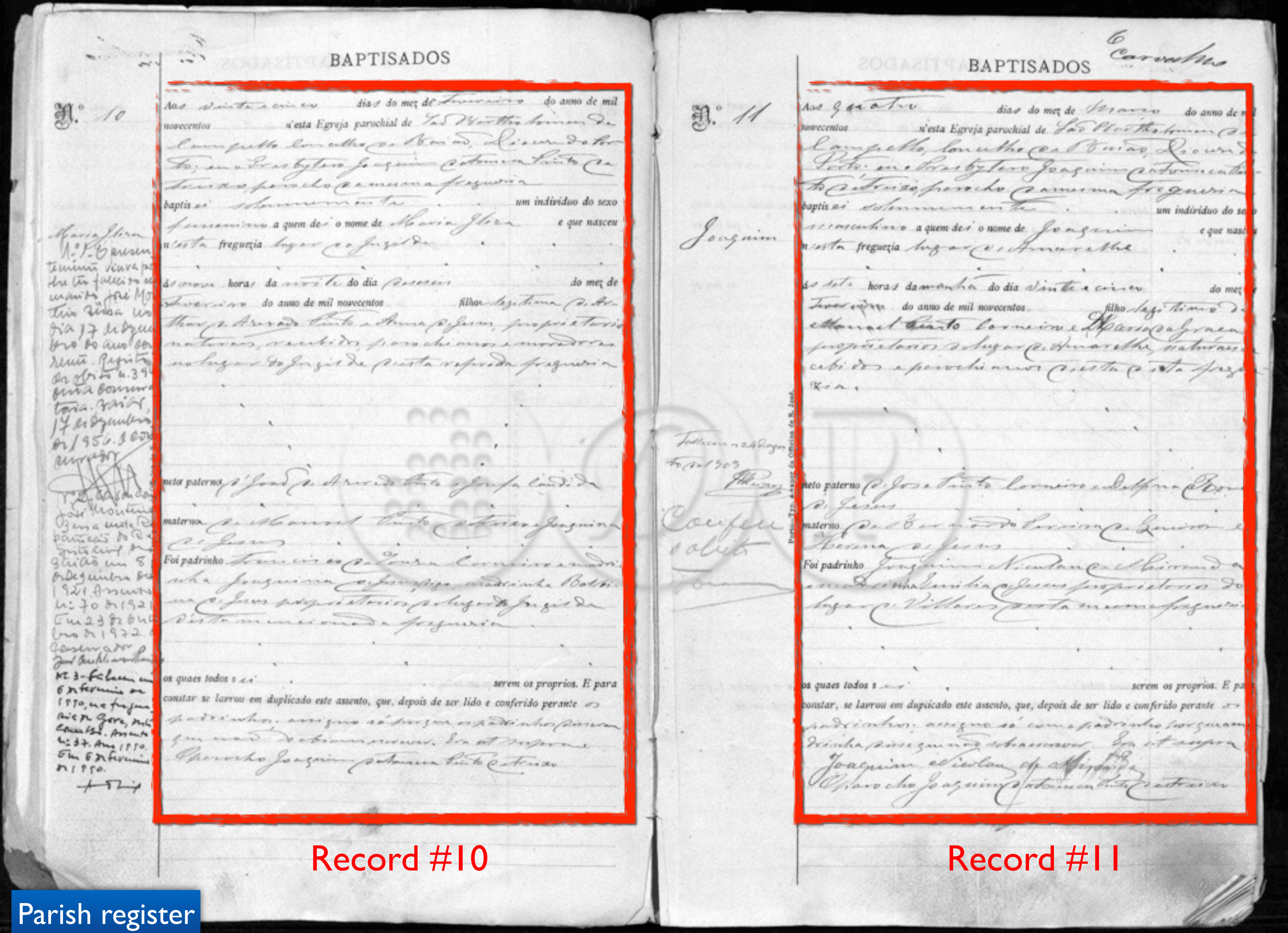
os quaes todos sei serem os proprios. E para constar, se lavrou em duplicado este assento, que, depois de ser lido e conferido perante os padrinhos, assigno só com o padrinho porque a madrinha disse que não sabia escrever. Era ut supra

Joaquim Nicolau de Miranda

O parocho Joaquim Antonio Pinto Teixeira

Porto—Typ. a vapor da Officina de S. José.

Parish register



BAPTISADOS

N.º 10

Aos vinte e cinco dias do mez de Fevereiro do anno de mil novecentos n'esta Egreja parochial de São Bartholomeu de Campello, concelho de Baião, diocese do Porto, eu o Presbytero Joaquim Antonio Pinto de Freitas, parocho da mesma freguezia baptisei solemnemente um individuo do sexo feminino a quem dei o nome de Maria Gloria e que nasceu n'esta freguezia lugar do Jugido

ás nove horas da noite do dia dezeseis do mez de Fevereiro do anno de mil novecentos filho legitima de Arthur de Azevedo Pinto e Anna de Jesus, proprietarios naturaes, recebidos, parochianos e moradores no lugar do Jugido d'esta referida freguezia

neta paterna de João de Azevedo Pinto e Josepha Candida

materna de Manoel Pinto de Freitas e Joaquina de Jesus

Foi padrinho Francisco da Silva Carneiro e madrinha Joaquina de Jesus, digo, madrinha Bellina de Jesus proprietarios do lugar do Jugido d'esta mencionada freguezia

os quaes todos sei serem os proprios. E para constar se lavrou em duplicado este assento, que, depois de ser lido e conferido perante os padrinhos, assigno só porque os padrinhos disseram que não sabiam escrever. Era ut supra.

O parocho Joaquim Antonio Pinto de Freitas

Maria Gloria

N.º 1 - O presente tem inum viuva por lhe ter fallecido o marido José Monteiro Sousa no dia 17 de Dezembro do ano corrente. Registo de obito n.º 394 desta Conservatoria. Baião, 17 de Dezembro de 1956. O Conservador

N.º 2 - Casou com José Monteiro Sousa na Repartição do Registo Civil de Baião em 8 de Dezembro de 1921. Assento n.º 70 de 1921. Em 23 de Outubro de 1972. O Conservador

N.º 3 - Faleceu em 6 de Fevereiro de 1950, na freguesia de Gove, concelho de Baião. Assento n.º 37. Ano 1950. Em 6 de Fevereiro de 1950.

Record #10

6 Carvalho

BAPTISADOS

N.º 11

Aos quatro dias do mez de Março do anno de mil novecentos n'esta Egreja parochial de São Bartholomeu de Campello, concelho de Baião, diocese do Porto, eu o Presbytero Joaquim Antonio Pinto de Freitas, parocho da mesma freguezia baptisei solemnemente um individuo do sexo masculino a quem dei o nome de Joaquim e que nasceu n'esta freguezia lugar de Amarelhe

ás sete horas da manhã do dia vinte e cinco do mez de Fevereiro do anno de mil novecentos filho legitimo de Manoel Pinto Carneiro e Maria da Graça proprietarios do lugar de Amarelhe, naturaes, recebidos e parochianos d'esta dita freguezia.

neto paterno de José Pinto Carneiro e Delphina Rosa de Jesus

materno de Bernardo Pereira de Souza e Thereza de Jesus

Foi padrinho Joaquim Nicolau de Miranda e madrinha Emilia de Jesus proprietarios do lugar de Villares d'esta mesma freguezia

os quaes todos sei serem os proprios. E para constar, se lavrou em duplicado este assento, que, depois de ser lido e conferido perante os padrinhos, assigno só com o padrinho porque a madrinha disse que não sabia escrever. Era ut supra.

Joaquim Nicolau de Miranda

O parocho Joaquim Antonio Pinto de Freitas

Joaquim

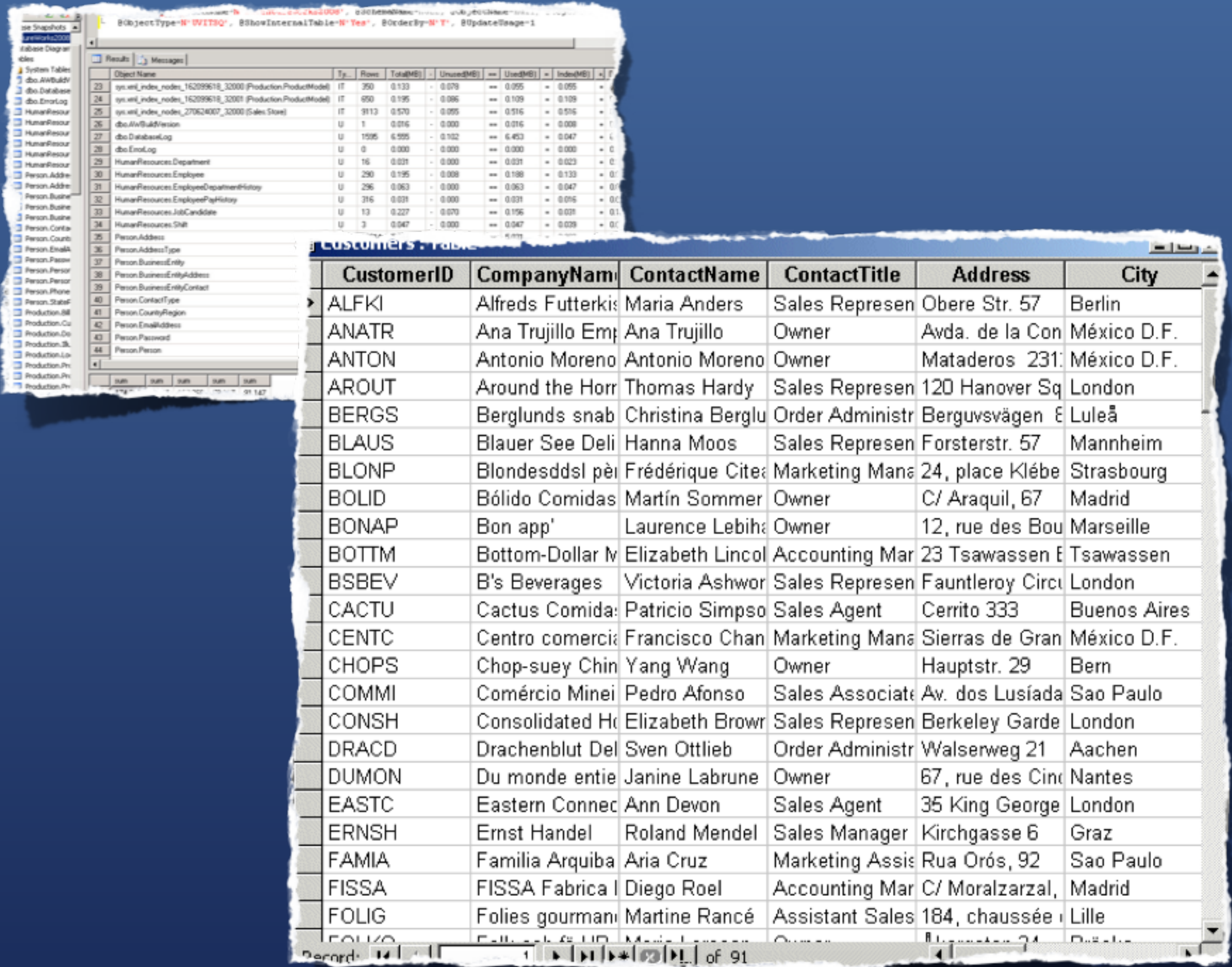
Falleceu a 24 d'Agosto de 1903

Conferi o obito

Porto—Typ. a vapor da Officina de S. José

Record #11

Parish register

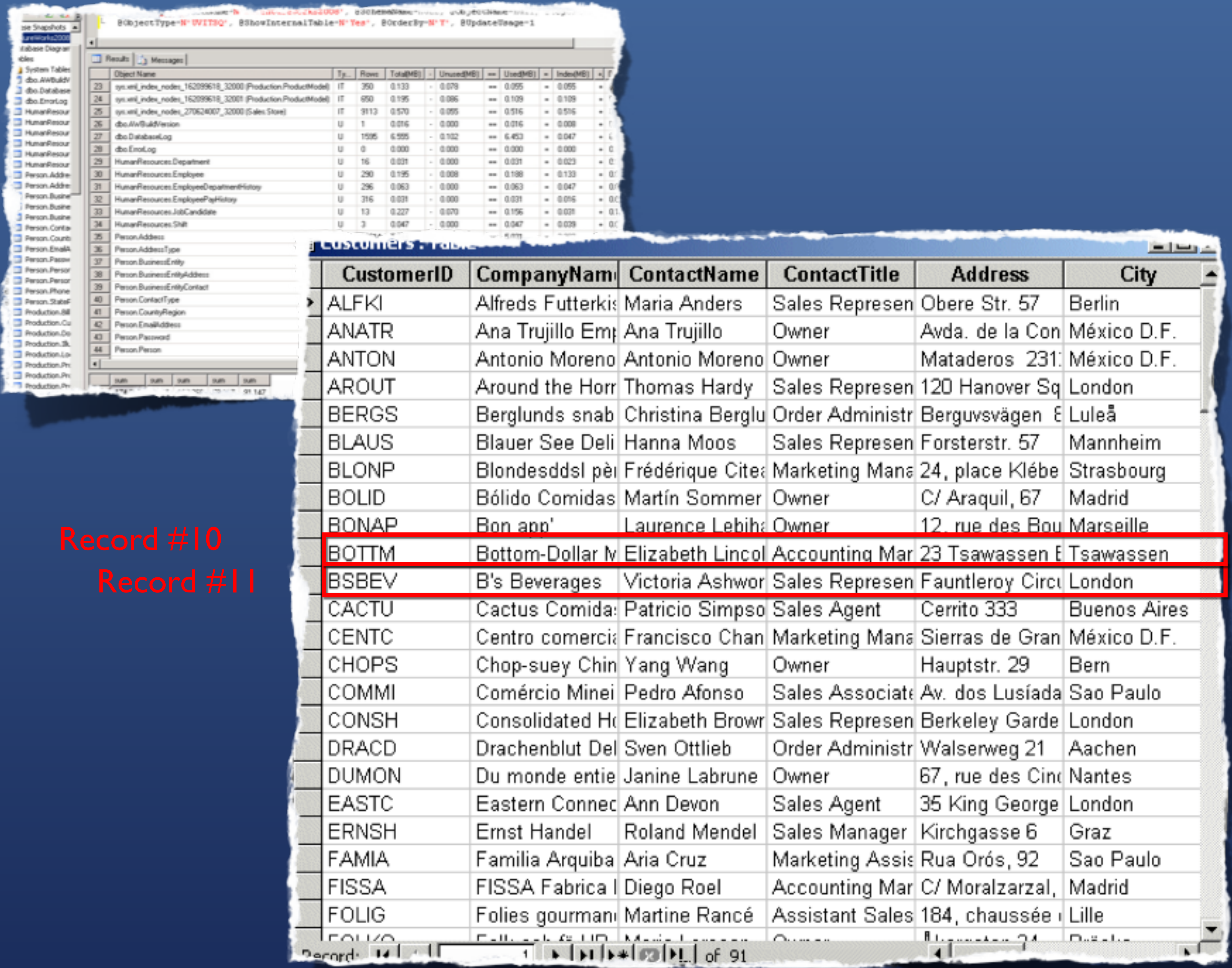


```
@ObjectType=N'UVITSQ', @ShowInternalTable=N'Yes', @OrderBy=N'T', @UpdateUsage=1
```

	Object Name	Ty...	Rows	Total(MB)	-	Unused(MB)	==	Used(MB)	=	Index(MB)	+
23	sys.xml_index_nodes_162099618_32000 (Production.ProductModel)	IT	350	0.133	-	0.078	==	0.055	=	0.055	+
24	sys.xml_index_nodes_162099618_32001 (Production.ProductModel)	IT	650	0.195	-	0.086	==	0.109	=	0.109	+
25	sys.xml_index_nodes_270624007_32000 (Sales.Store)	IT	9113	0.570	-	0.055	==	0.516	=	0.516	+
26	dbo.AWBuildVersion	U	1	0.016	-	0.000	==	0.016	=	0.008	+
27	dbo.DatabaseLog	U	1595	6.555	-	0.102	==	6.453	=	0.047	+
28	dbo.ErrorLog	U	0	0.000	-	0.000	==	0.000	=	0.000	+
29	HumanResources.Department	U	16	0.031	-	0.000	==	0.031	=	0.023	+
30	HumanResources.Employee	U	290	0.195	-	0.008	==	0.188	=	0.133	+
31	HumanResources.EmployeeDepartmentHistory	U	296	0.063	-	0.000	==	0.063	=	0.047	+
32	HumanResources.EmployeePayHistory	U	316	0.031	-	0.000	==	0.031	=	0.016	+
33	HumanResources.JobCandidate	U	13	0.227	-	0.070	==	0.156	=	0.031	+
34	HumanResources.Shift	U	3	0.047	-	0.000	==	0.047	=	0.039	+
35	Person.Address										
36	Person.AddressType										
37	Person.BusinessEntity										
38	Person.BusinessEntityAddress										
39	Person.BusinessEntityContact										
40	Person.ContactType										
41	Person.CountryRegion										
42	Person.EmailAddress										
43	Person.Password										
44	Person.Person										
	sum	sum	sum	sum	sum						
	[illegible]		[illegible]	[illegible]	91.147						

CustomerID	CompanyNam	ContactName	ContactTitle	Address	City
ALFKI	Alfreds Futterkis	Maria Anders	Sales Represen	Obere Str. 57	Berlin
ANATR	Ana Trujillo Emp	Ana Trujillo	Owner	Avda. de la Con	México D.F.
ANTON	Antonio Moreno	Antonio Moreno	Owner	Mataderos 231	México D.F.
AROUT	Around the Horr	Thomas Hardy	Sales Represen	120 Hanover Sq	London
BERGS	Berglunds snab	Christina Berglu	Order Administr	Berguvsvägen 8	Luleå
BLAUS	Blauer See Deli	Hanna Moos	Sales Represen	Forsterstr. 57	Mannheim
BLONP	Blondesddsl pèr	Frédérique Cite	Marketing Mana	24, place Klébe	Strasbourg
BOLID	Bólido Comidas	Martín Sommer	Owner	C/ Araquil, 67	Madrid
BONAP	Bon app'	Laurence Lebiha	Owner	12, rue des Bou	Marseille
BOTTM	Bottom-Dollar M	Elizabeth Lincol	Accounting Mar	23 Tsawassen B	Tsawassen
BSBEV	B's Beverages	Victoria Ashwor	Sales Represen	Fauntleroy Circu	London
CACTU	Cactus Comida	Patricio Simpso	Sales Agent	Cerrito 333	Buenos Aires
CENTC	Centro comercia	Francisco Chan	Marketing Mana	Sierras de Gran	México D.F.
CHOPS	Chop-suey Chin	Yang Wang	Owner	Hauptstr. 29	Bern
COMMI	Comércio Minei	Pedro Afonso	Sales Associate	Av. dos Lusíada	Sao Paulo
CONSH	Consolidated H	Elizabeth Brown	Sales Represen	Berkeley Garde	London
DRACD	Drachenblut Del	Sven Ottlieb	Order Administr	Walserweg 21	Aachen
DUMON	Du monde entie	Janine Labrune	Owner	67, rue des Cinc	Nantes
EASTC	Eastern Connec	Ann Devon	Sales Agent	35 King George	London
ERNSH	Ernst Handel	Roland Mendel	Sales Manager	Kirchgasse 6	Graz
FAMIA	Familia Arquiba	Aria Cruz	Marketing Assis	Rua Orós, 92	Sao Paulo
FISSA	FISSA Fabrica I	Diego Roel	Accounting Mar	C/ Moralzarzal,	Madrid
FOLIG	Folies gourman	Martine Rancé	Assistant Sales	184, chaussée	Lille
FOLKO	Folk och fä HB	Maria Larsson	Owner	Åkergatan 24	Bräcke

Record: 1 of 91



```
@ObjectType=N'UVITSQ', @ShowInternalTable=N'Yes', @OrderBy=N'T', @UpdateUsage=1
```

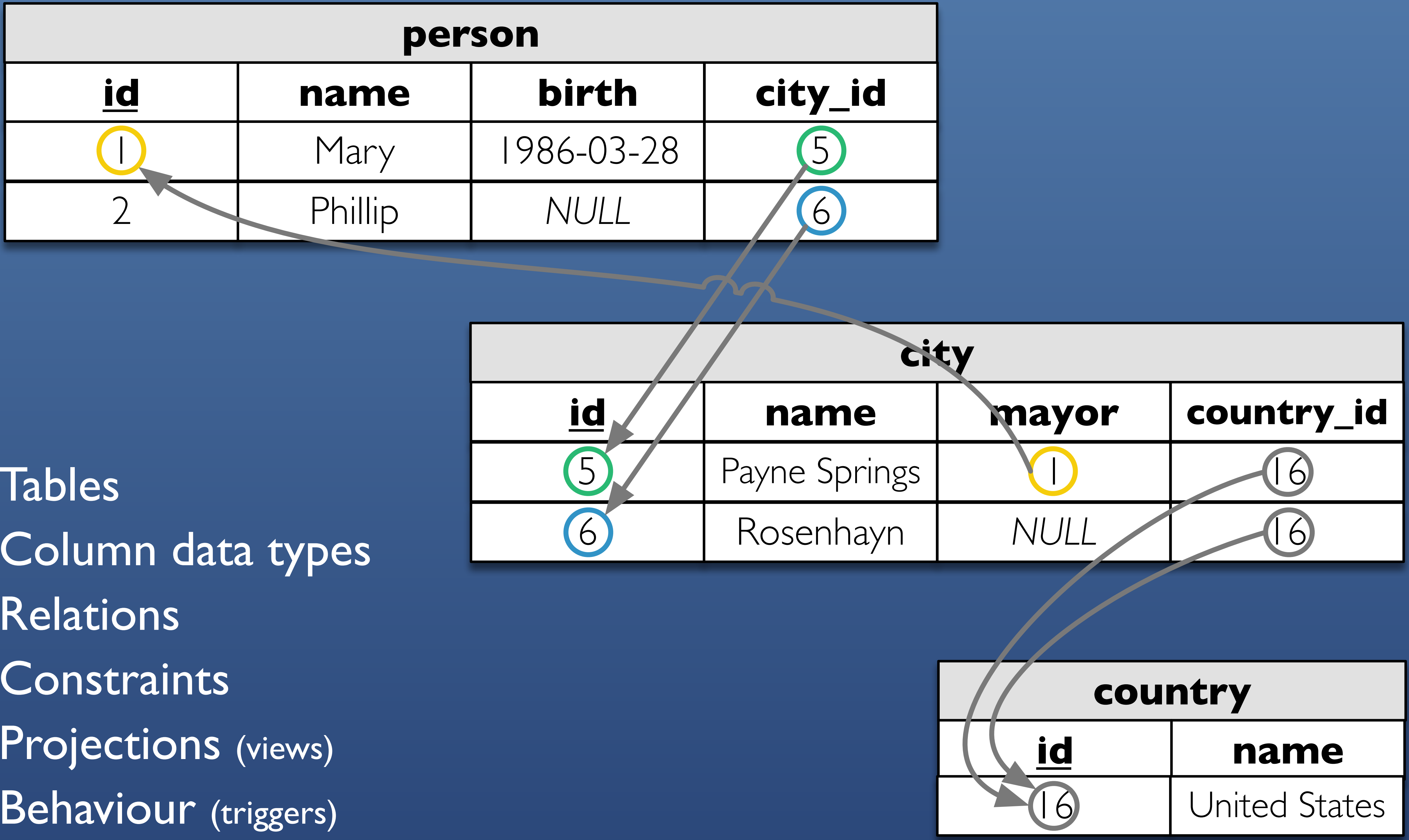
	Object Name	Ty...	Rows	Total(MB)	-	Unused(MB)	==	Used(MB)	=	Index(MB)
23	sys.xml_index_nodes_1620996l8_32000 (Production.ProductModel)	IT	350	0.133	-	0.078	==	0.055	=	0.055
24	sys.xml_index_nodes_1620996l8_32001 (Production.ProductModel)	IT	650	0.195	-	0.086	==	0.109	=	0.109
25	sys.xml_index_nodes_270624007_32000 (Sales.Store)	IT	9113	0.570	-	0.055	==	0.516	=	0.516
26	dbo.AWBuildVersion	U	1	0.016	-	0.000	==	0.016	=	0.008
27	dbo.DatabaseLog	U	1595	6.555	-	0.102	==	6.453	=	0.047
28	dbo.ErrorLog	U	0	0.000	-	0.000	==	0.000	=	0.000
29	HumanResources.Department	U	16	0.031	-	0.000	==	0.031	=	0.023
30	HumanResources.Employee	U	290	0.195	-	0.008	==	0.188	=	0.133
31	HumanResources.EmployeeDepartmentHistory	U	296	0.063	-	0.000	==	0.063	=	0.047
32	HumanResources.EmployeePayHistory	U	316	0.031	-	0.000	==	0.031	=	0.016
33	HumanResources.JobCandidate	U	13	0.227	-	0.070	==	0.156	=	0.031
34	HumanResources.Shift	U	3	0.047	-	0.000	==	0.047	=	0.039
35	Person.Address									
36	Person.AddressType									
37	Person.BusinessEntity									
38	Person.BusinessEntityAddress									
39	Person.BusinessEntityContact									
40	Person.ContactType									
41	Person.CountryRegion									
42	Person.EmailAddress									
43	Person.Password									
44	Person.Person									

CustomerID	CompanyNam	ContactName	ContactTitle	Address	City
ALFKI	Alfreds Futterkis	Maria Anders	Sales Represen	Obere Str. 57	Berlin
ANATR	Ana Trujillo Emp	Ana Trujillo	Owner	Avda. de la Con	México D.F.
ANTON	Antonio Moreno	Antonio Moreno	Owner	Mataderos 231	México D.F.
AROUT	Around the Horn	Thomas Hardy	Sales Represen	120 Hanover Sq	London
BERGS	Berglunds snab	Christina Berglu	Order Administr	Berguvsvägen 8	Luleå
BLAUS	Blauer See Deli	Hanna Moos	Sales Represen	Forsterstr. 57	Mannheim
BLONP	Blondesddsl pèi	Frédérique Cite	Marketing Mana	24, place Klébe	Strasbourg
BOLID	Bólido Comidas	Martín Sommer	Owner	C/ Araquil, 67	Madrid
BONAP	Bon app'	Laurence Lebiha	Owner	12, rue des Bou	Marseille
BOTTM	Bottom-Dollar M	Elizabeth Lincol	Accounting Mar	23 Tsawassen E	Tsawassen
BSBEV	B's Beverages	Victoria Ashwor	Sales Represen	Fauntleroy Circu	London
CACTU	Cactus Comida	Patricio Simpso	Sales Agent	Cerrito 333	Buenos Aires
CENTC	Centro comercia	Francisco Chan	Marketing Mana	Sierras de Gran	México D.F.
CHOPS	Chop-suey Chin	Yang Wang	Owner	Hauptstr. 29	Bern
COMMI	Comércio Minei	Pedro Afonso	Sales Associate	Av. dos Lusíada	Sao Paulo
CONSH	Consolidated Ho	Elizabeth Brown	Sales Represen	Berkeley Garde	London
DRACD	Drachenblut Del	Sven Ottlieb	Order Administr	Walserweg 21	Aachen
DUMON	Du monde entie	Janine Labrune	Owner	67, rue des Cinq	Nantes
EASTC	Eastern Connec	Ann Devon	Sales Agent	35 King George	London
ERNSH	Ernst Handel	Roland Mendel	Sales Manager	Kirchgasse 6	Graz
FAMIA	Familia Arquiba	Aria Cruz	Marketing Assis	Rua Orós, 92	Sao Paulo
FISSA	FISSA Fabrica I	Diego Roel	Accounting Mar	C/ Moralzarzal,	Madrid
FOLIG	Folies gourman	Martine Rancé	Assistant Sales	184, chaussée	Lille
FOLKO	Folk och fä HB	Maria Larsson	Owner	Åkergatan 24	Bräcke

Record: |◀ ◀ 1 ▶ ▶| ▶* of 91

Record #10

Record #11



person

<u>id</u>	name	birth	city_id
1	Mary	1986-03-28	5
2	Phillip	*NULL*	6

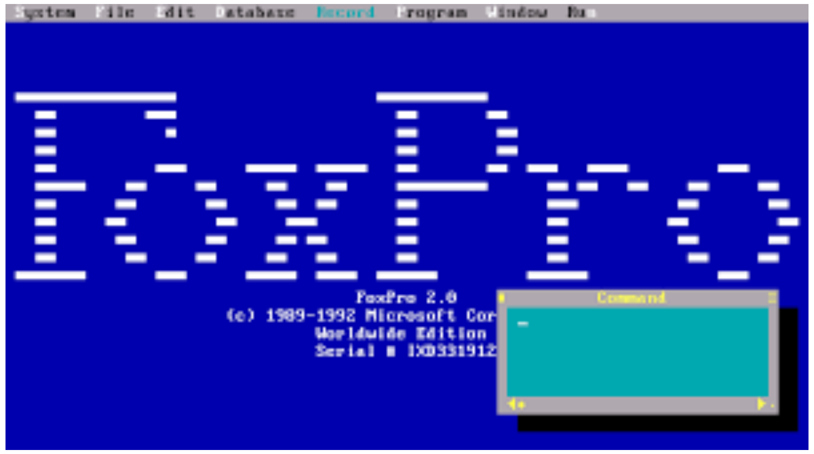
city

<u>id</u>	name	mayor	country_id
5	Payne Springs	1	16
6	Rosenhayn	*NULL*	16

country

<u>id</u>	name
16	United States

- Tables
- Column data types
- Relations
- Constraints
- Projections (views)
- Behaviour (triggers)
- Other (users, permissions, etc.)



```
FoxPro 2.0
(c) 1989-1992 Microsoft Cor
Worldwide Edition
Serial # 1XD331912
```

SQL:99

SIARD 2.0

Specification for archiving relational databases

Special thanks to Marcel Büchler from the Swiss Federal Archives (SFA) for some of the graphics and information included in this presentation.

It’s a vendor independent database archiving format

Stands for Software Independent Archiving of Relational Databases

Version 2 is a joint effort of the Swiss Federal Archives, the E-ARK project and the eCH (eGovernment standards in Switzerland)

Based on international standards

Unicode, SQL:1999, XML, XML Schema, URI

SIARD 2 is E-Government Standard (eCH-0165)

Replaced version 1.0 in 2016

What it does…

Preserves information — not layout or interaction

The application and the business logic are not preserved

Preserves data — not code

Stored procedures, functions and other code-like properties are stored but not preserved

Preserves tables and relations —not dynamic data

Views are not be preserved in SIARD, but they can be materialised

Some technical details...

Data is stored in a folder-like structure

Optionally, the folder can be compressed as ZIP for storage saving purposes

A header folder stores database metadata

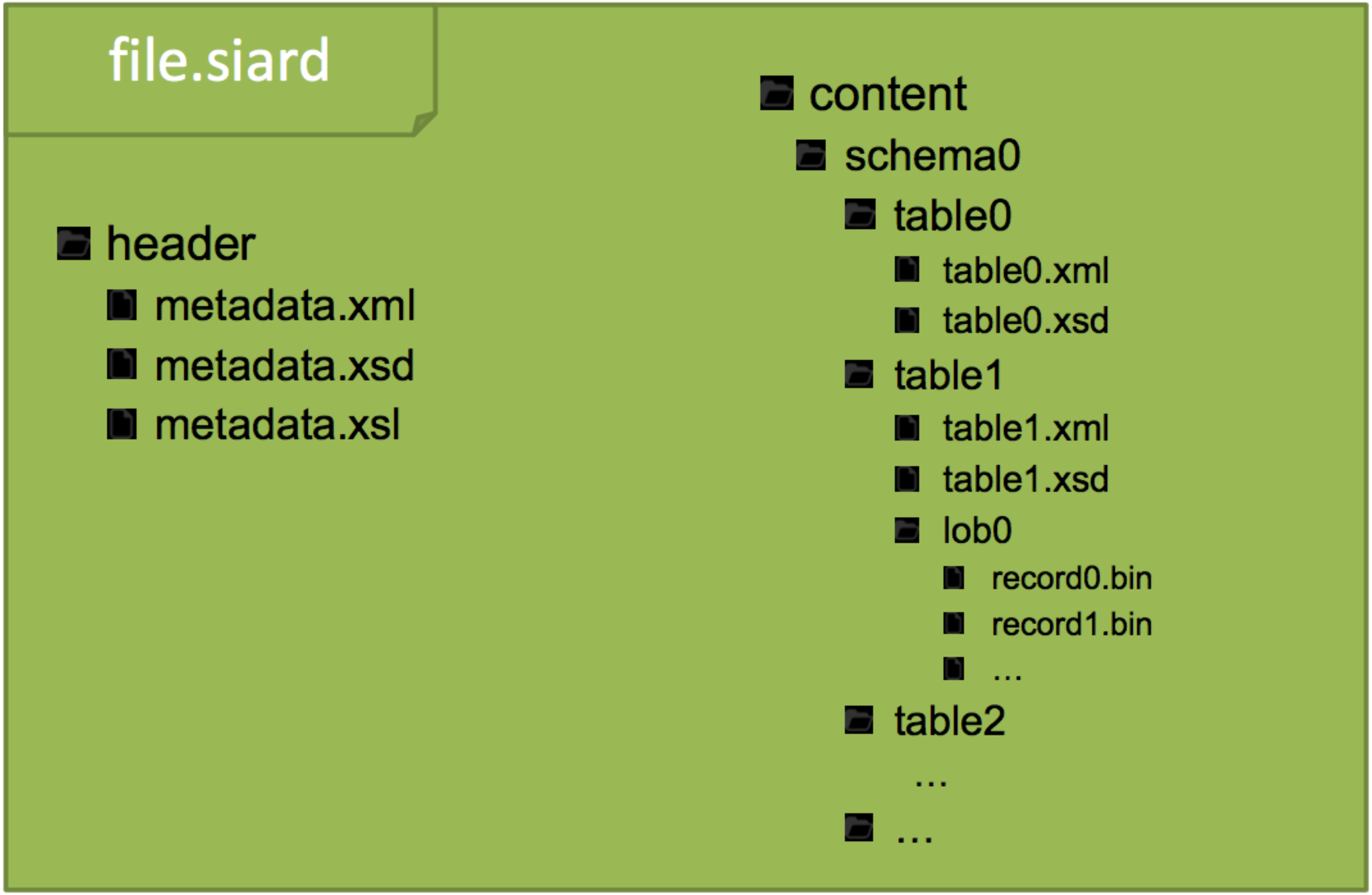
General information about the archived database and database structure stored as XML

A content folder stores database content

Tabular data stored as multiple XML files

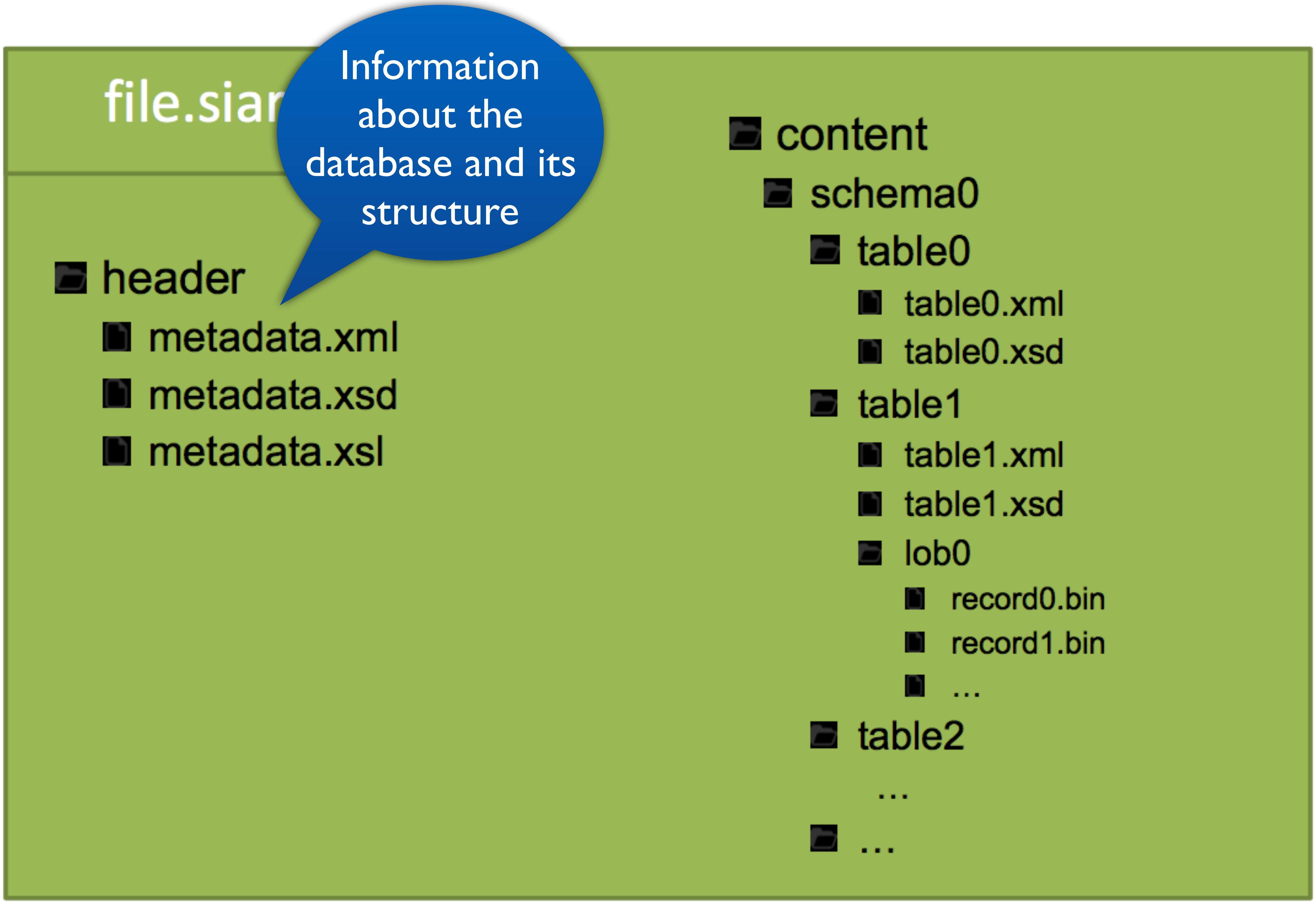
Binary objects are stored in 3 different ways

Inline, inside, outside



file.siard

- header
  - metadata.xml
  - metadata.xsd
  - metadata.xsl
- content
  - schema0
    - table0
      - table0.xml
      - table0.xsd
    - table1
      - table1.xml
      - table1.xsd
      - lob0
        - record0.bin
        - record1.bin
        - …
    - table2
    - …
    - …

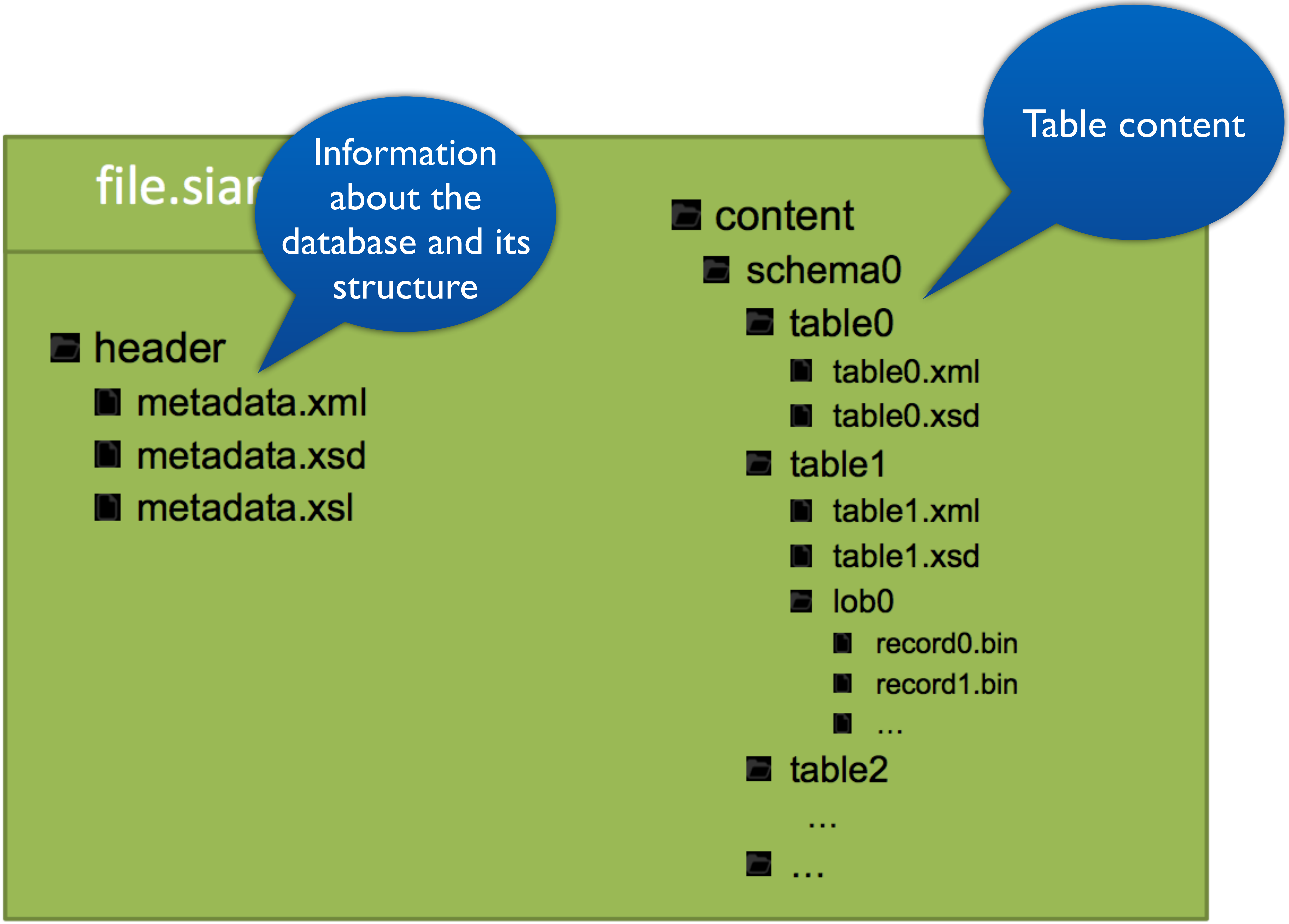


file.siar

Information about the database and its structure

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

- content
  - schema0
    - table0
      - table0.xml
      - table0.xsd
    - table1
      - table1.xml
      - table1.xsd
      - lob0
        - record0.bin
        - record1.bin
        - ...
    - table2
    - ...
  - ...



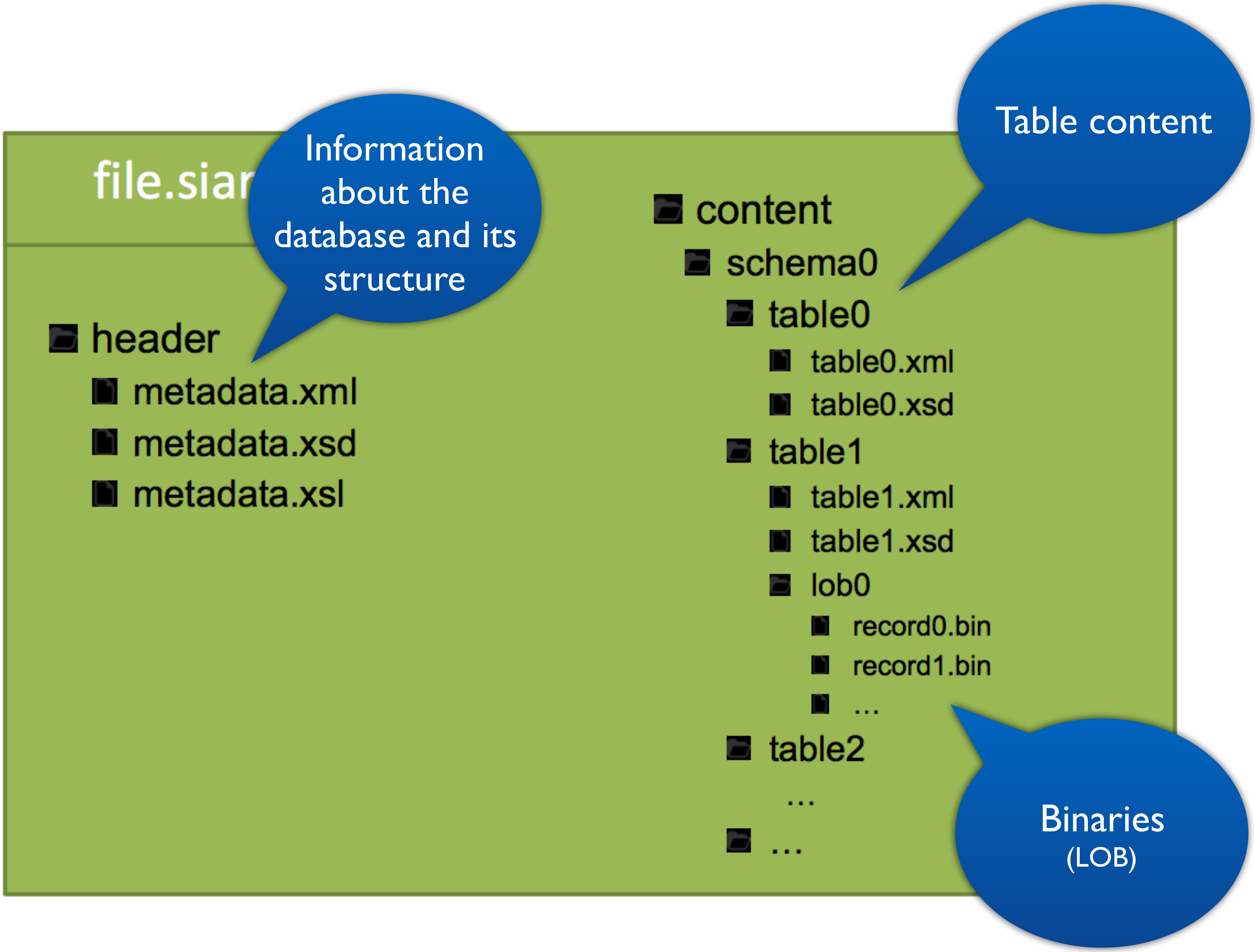
file.siar

Information about the database and its structure

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

Table content

- content
  - schema0
    - table0
      - table0.xml
      - table0.xsd
    - table1
      - table1.xml
      - table1.xsd
      - lob0
        - record0.bin
        - record1.bin
        - …
    - table2
      - …
    - …



file.siar

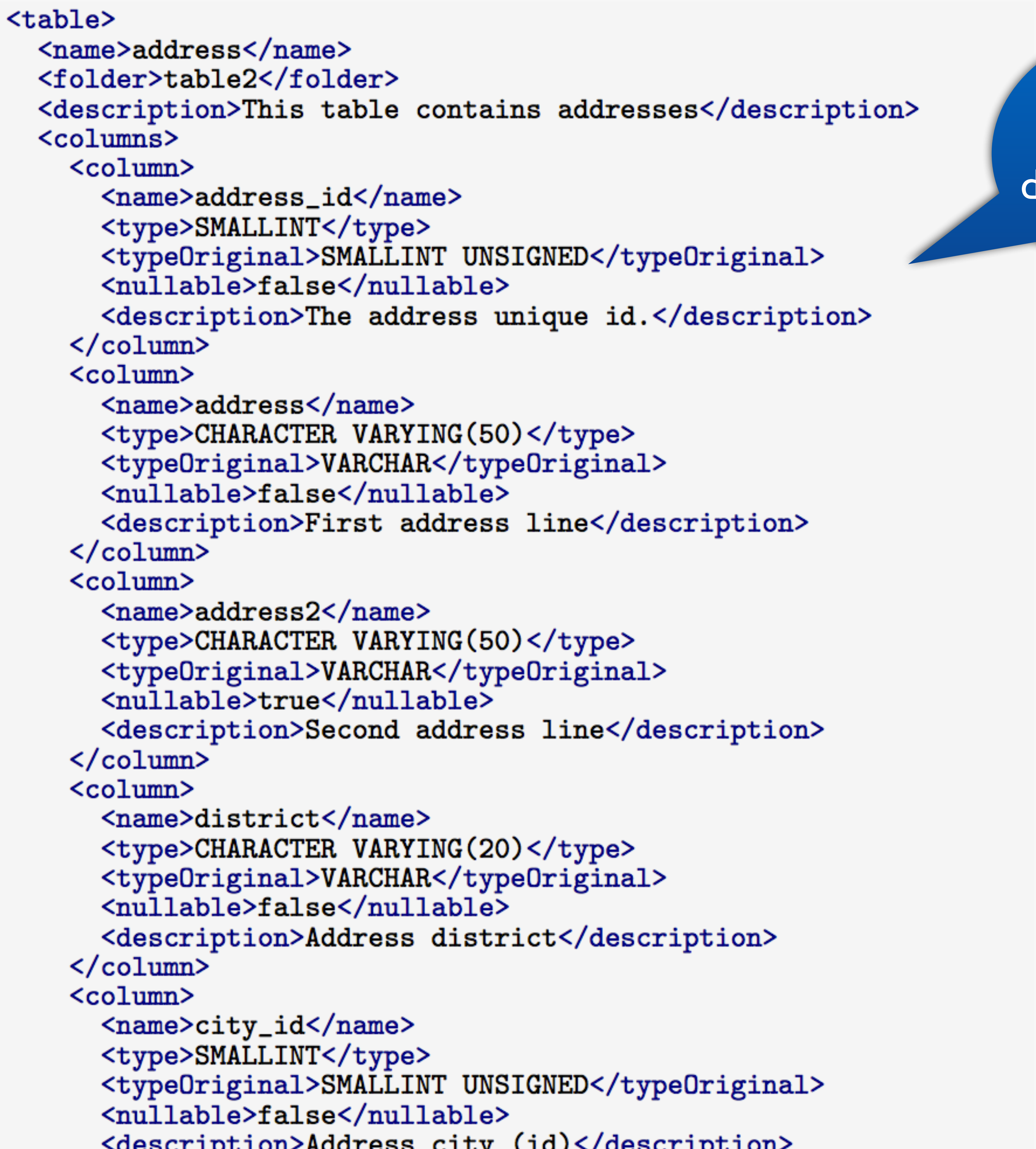
Information about the database and its structure

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

Table content

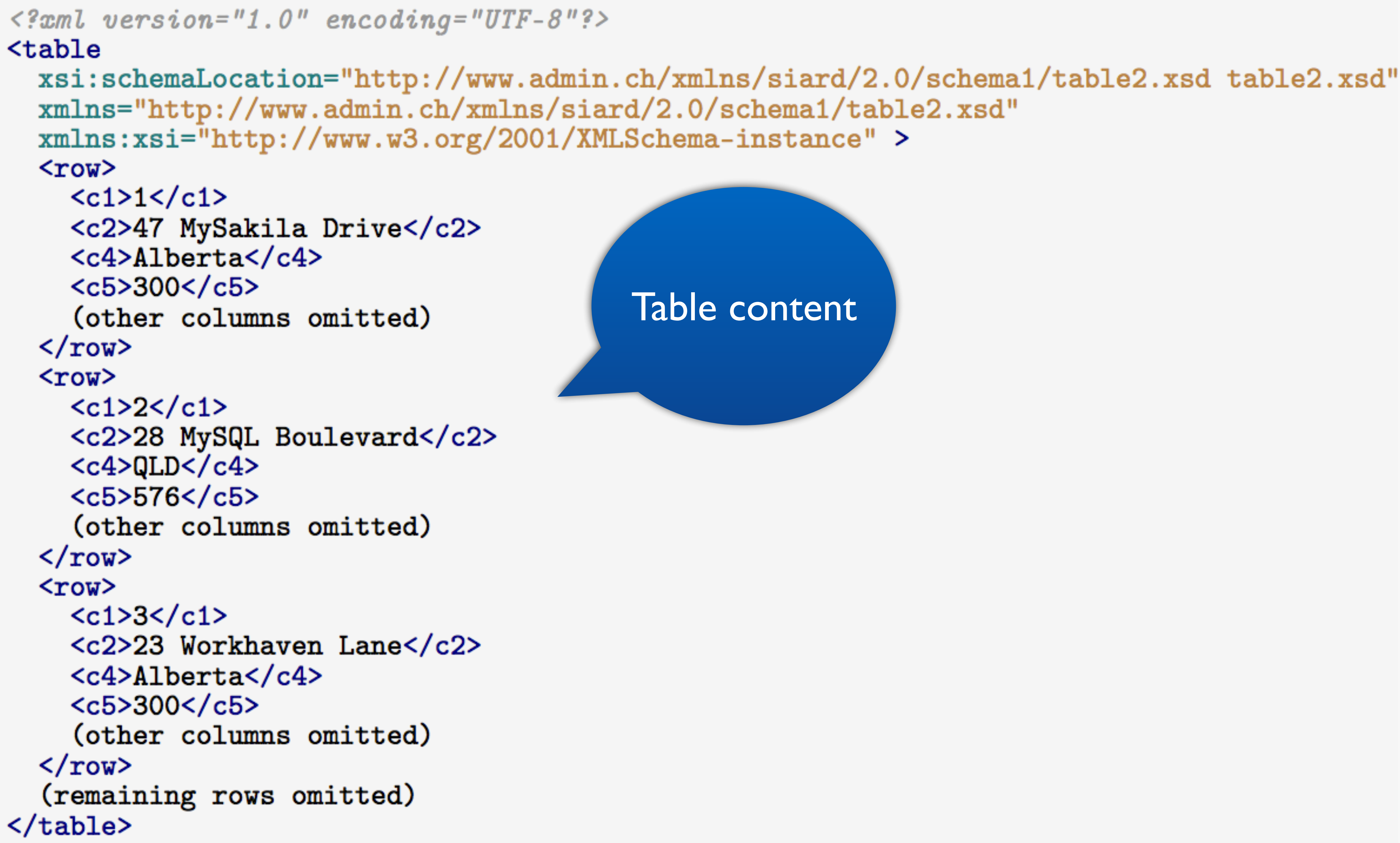
- content
  - schema0
    - table0
      - table0.xml
      - table0.xsd
    - table1
      - table1.xml
      - table1.xsd
      - lob0
        - record0.bin
        - record1.bin
        - …
    - table2
      …
    - …

Binaries
(LOB)



```
<table>
  <name>address</name>
  <folder>table2</folder>
  <description>This table contains addresses</description>
  <columns>
    <column>
      <name>address_id</name>
      <type>SMALLINT</type>
      <typeOriginal>SMALLINT UNSIGNED</typeOriginal>
      <nullable>false</nullable>
      <description>The address unique id.</description>
    </column>
    <column>
      <name>address</name>
      <type>CHARACTER VARYING(50)</type>
      <typeOriginal>VARCHAR</typeOriginal>
      <nullable>false</nullable>
      <description>First address line</description>
    </column>
    <column>
      <name>address2</name>
      <type>CHARACTER VARYING(50)</type>
      <typeOriginal>VARCHAR</typeOriginal>
      <nullable>true</nullable>
      <description>Second address line</description>
    </column>
    <column>
      <name>district</name>
      <type>CHARACTER VARYING(20)</type>
      <typeOriginal>VARCHAR</typeOriginal>
      <nullable>false</nullable>
      <description>Address district</description>
    </column>
    <column>
      <name>city_id</name>
      <type>SMALLINT</type>
      <typeOriginal>SMALLINT UNSIGNED</typeOriginal>
      <nullable>false</nullable>
      <description>Address city (id)</description>
```

Information about the database and its structure



```
<?xml version="1.0" encoding="UTF-8"?>
<table
  xsi:schemaLocation="http://www.admin.ch/xmlns/siard/2.0/schema1/table2.xsd table2.xsd"
  xmlns="http://www.admin.ch/xmlns/siard/2.0/schema1/table2.xsd"
  xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance" >
  <row>
    <c1>1</c1>
    <c2>47 MySakila Drive</c2>
    <c4>Alberta</c4>
    <c5>300</c5>
    (other columns omitted)
  </row>
  <row>
    <c1>2</c1>
    <c2>28 MySQL Boulevard</c2>
    <c4>QLD</c4>
    <c5>576</c5>
    (other columns omitted)
  </row>
  <row>
    <c1>3</c1>
    <c2>23 Workhaven Lane</c2>
    <c4>Alberta</c4>
    <c5>300</c5>
    (other columns omitted)
  </row>
  (remaining rows omitted)
</table>
```

Table content

Toolset

database preservation . toolkit

RODA

database visualization . toolkit

database preservation toolkit

Data extraction

RODA

database visualization toolkit

database preservation toolkit	Data extraction
RODA	Data archiving
database visualization toolkit	

database preservation toolkit	Data extraction
RODA	Data archiving
database visualization toolkit	Data visualisation

Use case scenario

An information system has been in use for 15 years

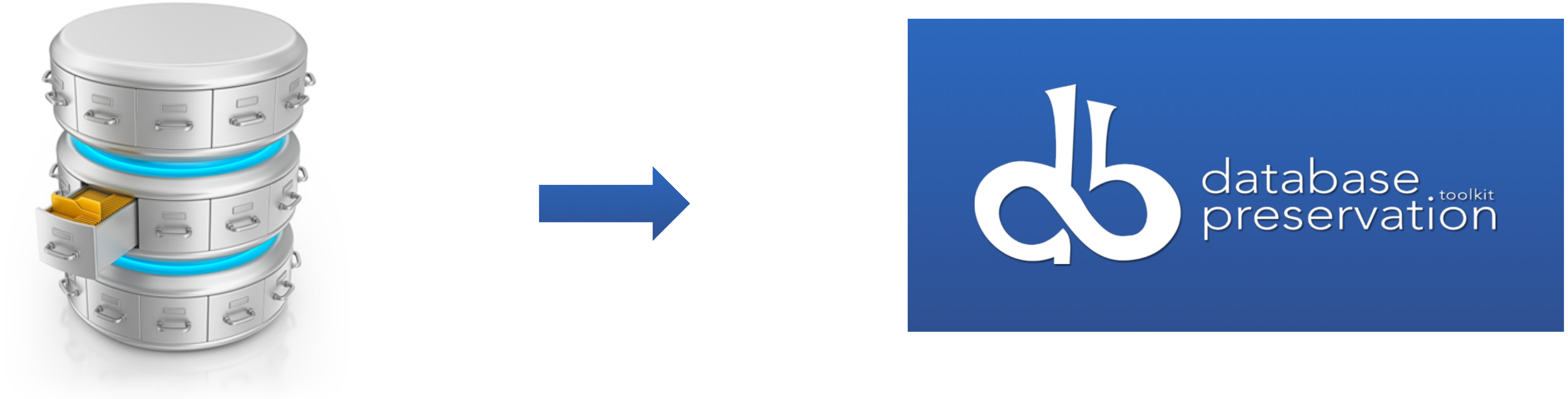
It is about to be decommissioned as it has been replaced by a new, more advanced system

There is an interest in maintaining the data produced by the legacy information system for legal and historical reasons

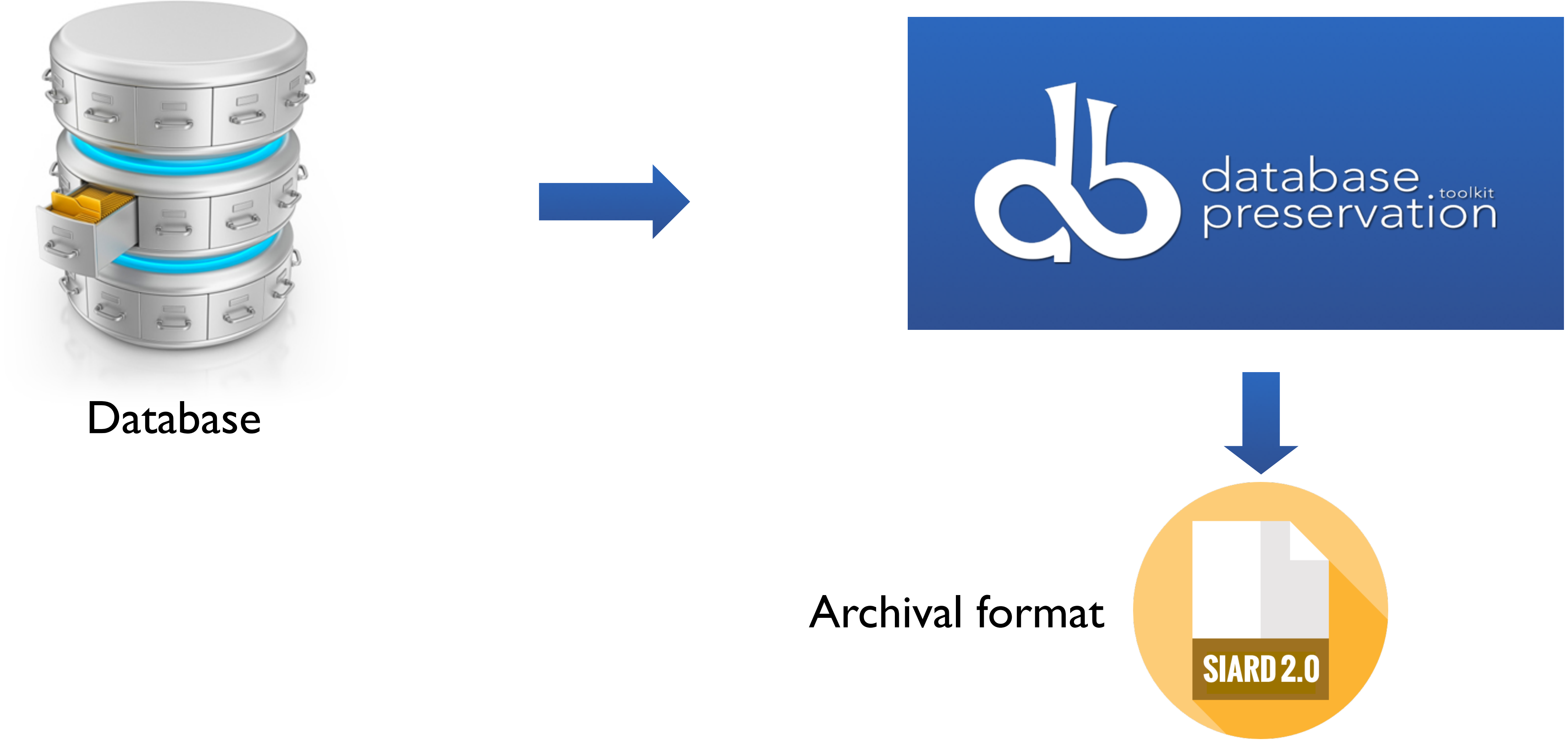
This means that the system’s database has been selected for long-term archival

Let’s see how it works…

Database

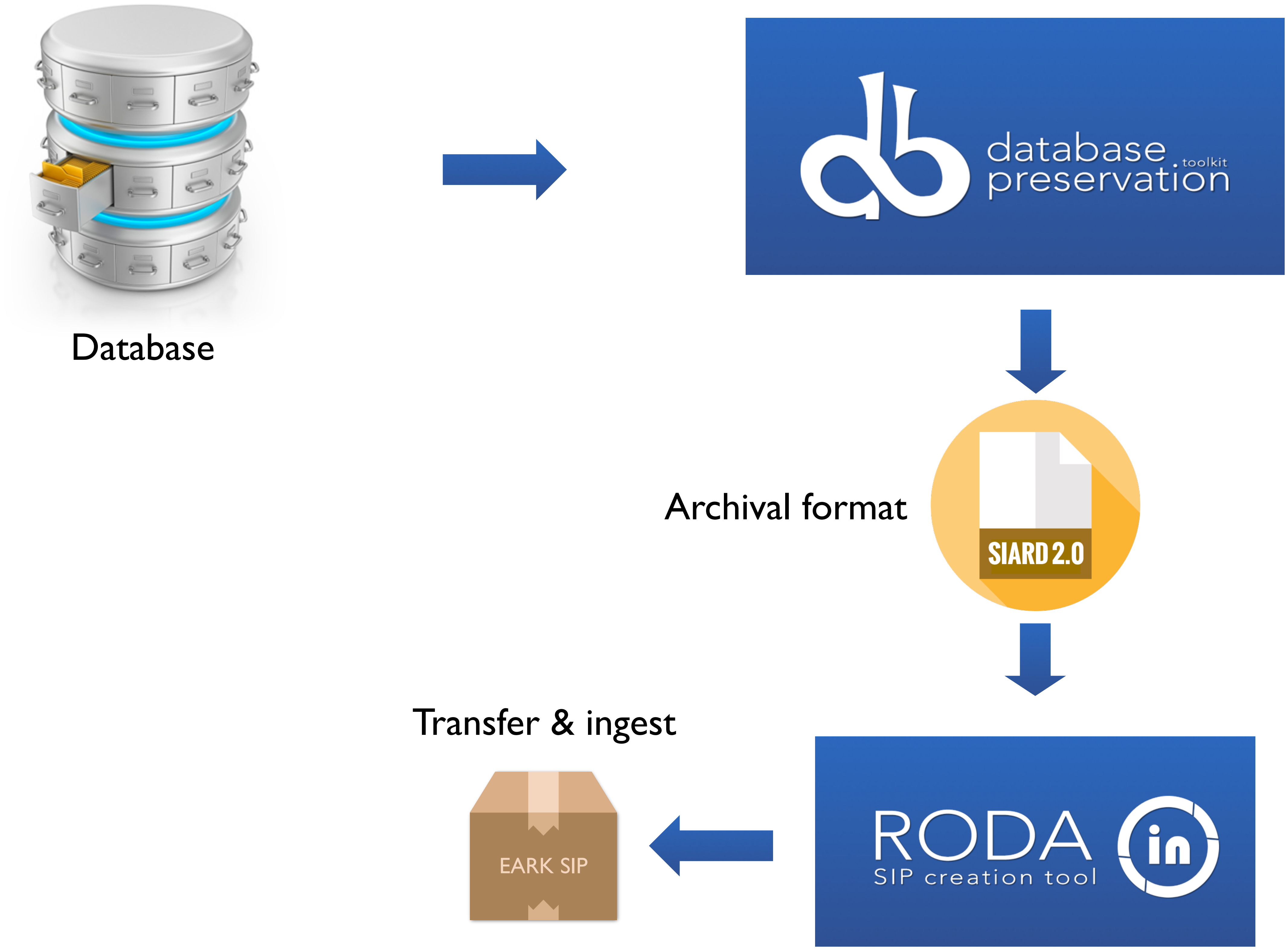


Database



Database

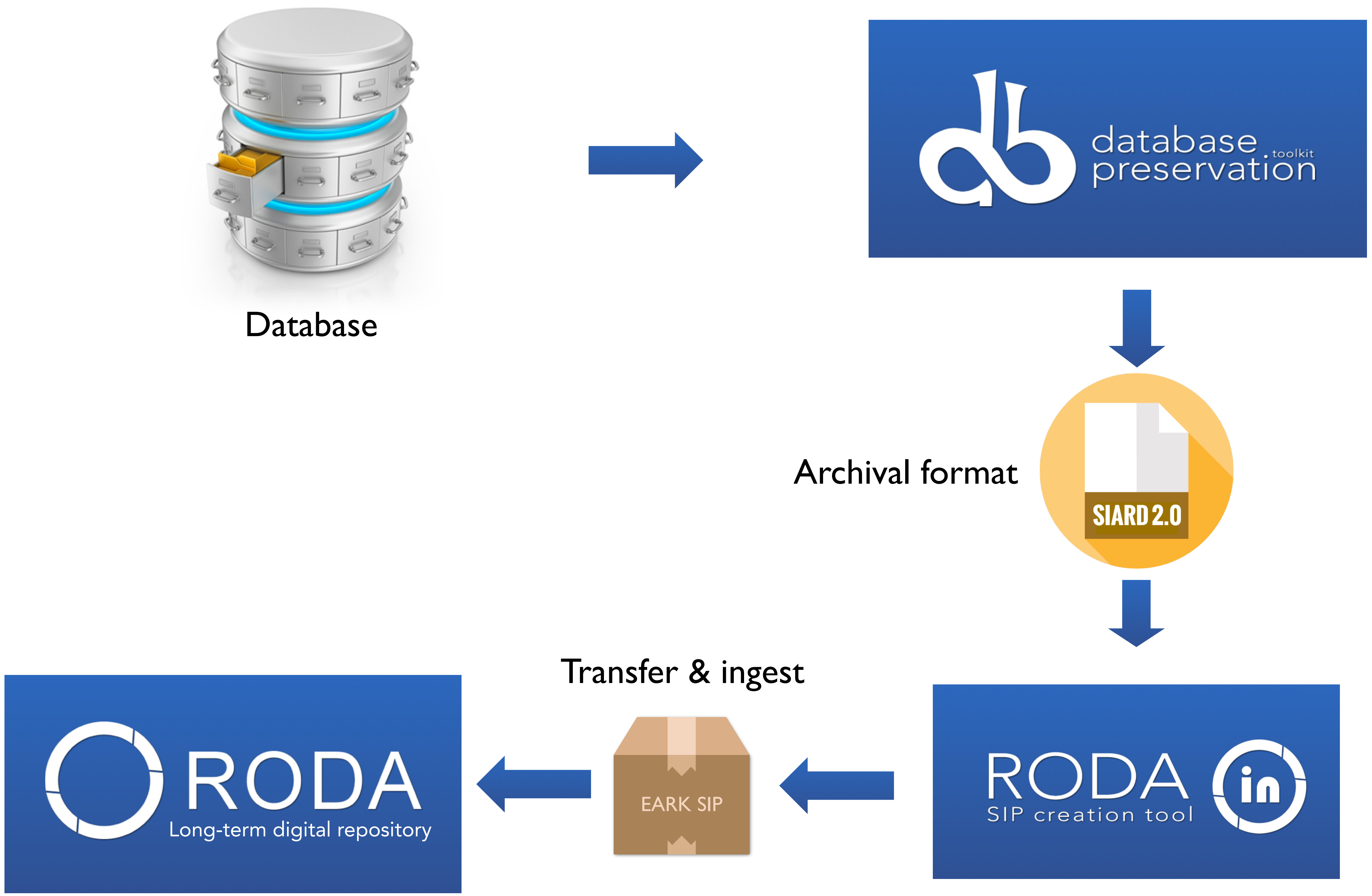
Archival format



Database

Archival format

Transfer & ingest

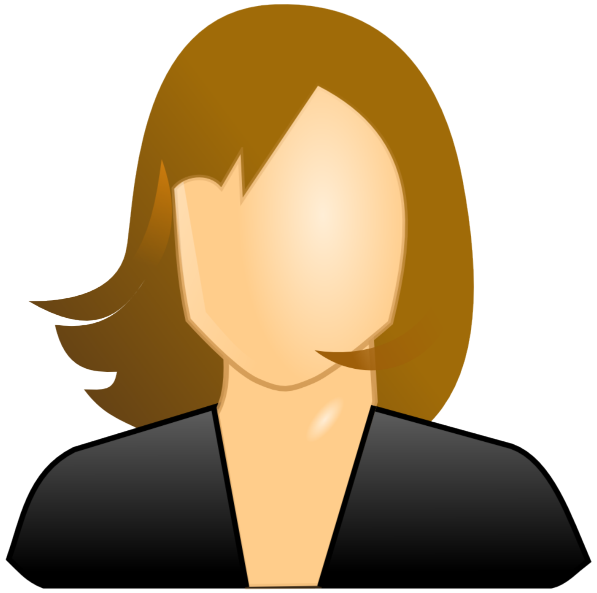


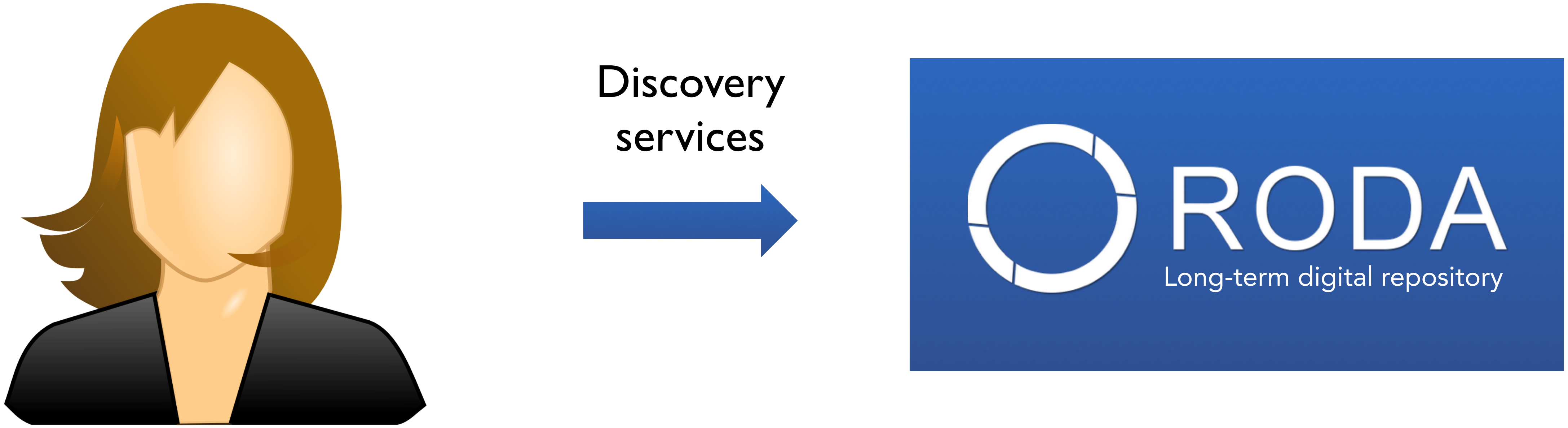
Database

Archival format

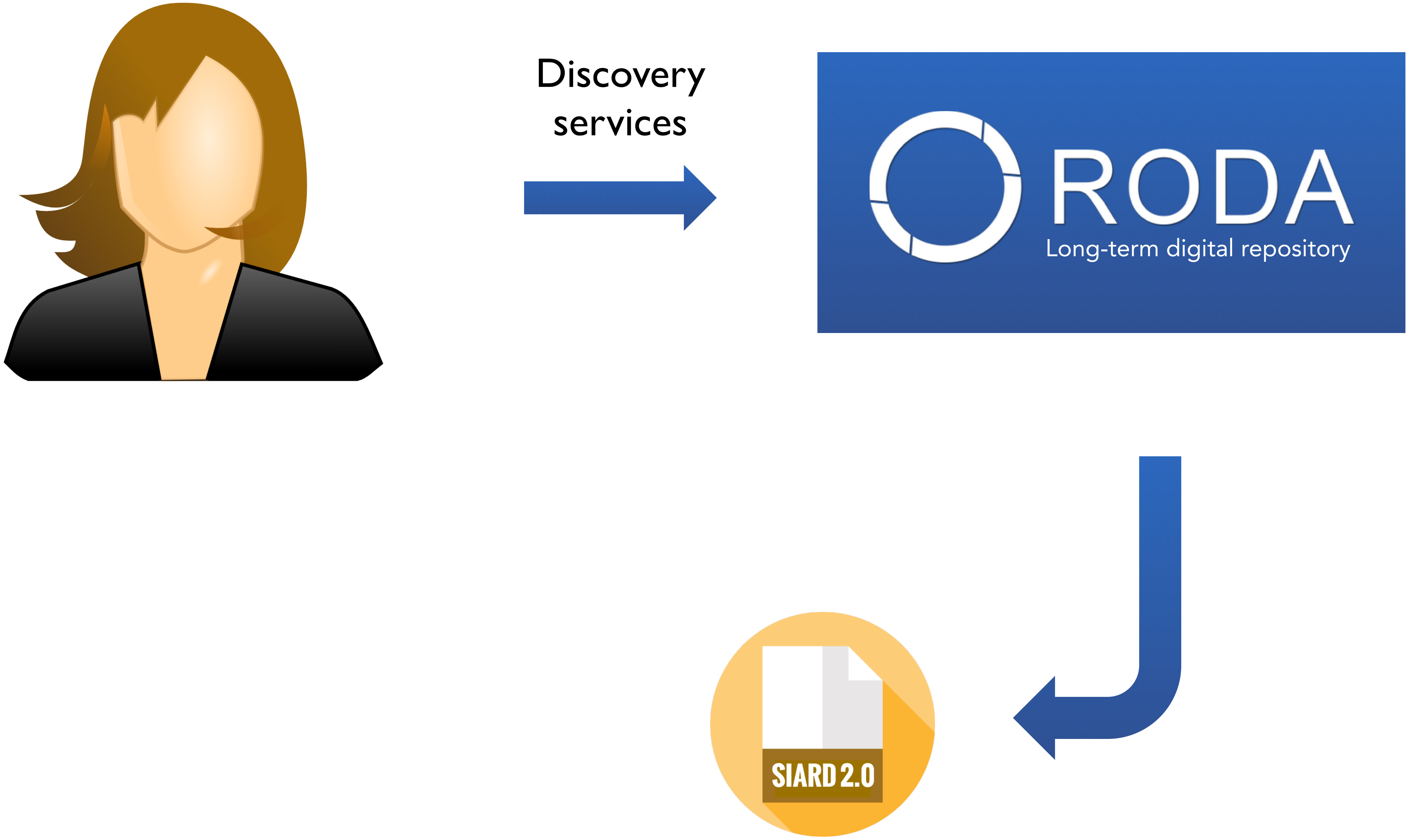
Transfer & ingest

Years later, a user wants to access the data…

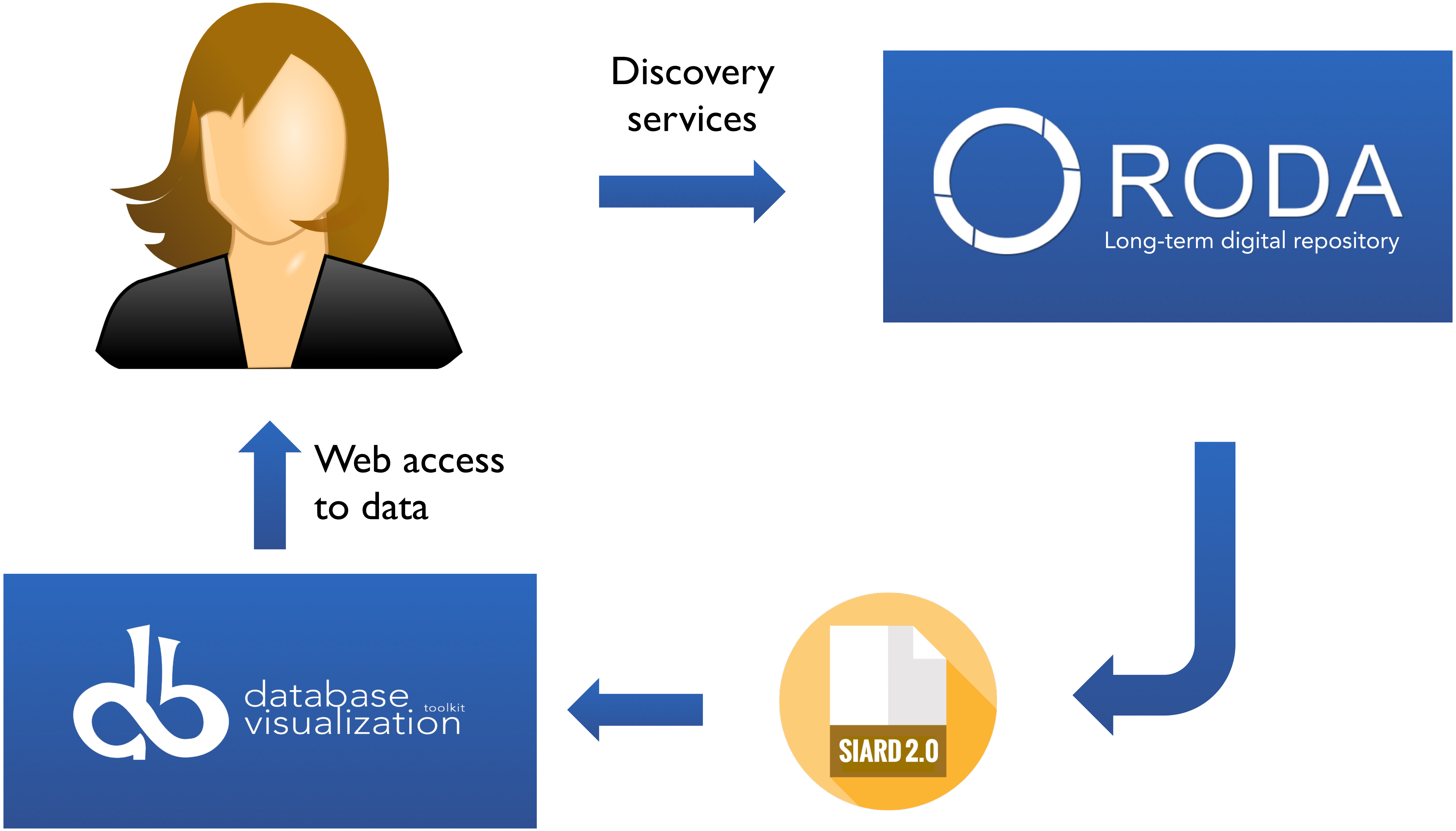




Discovery services

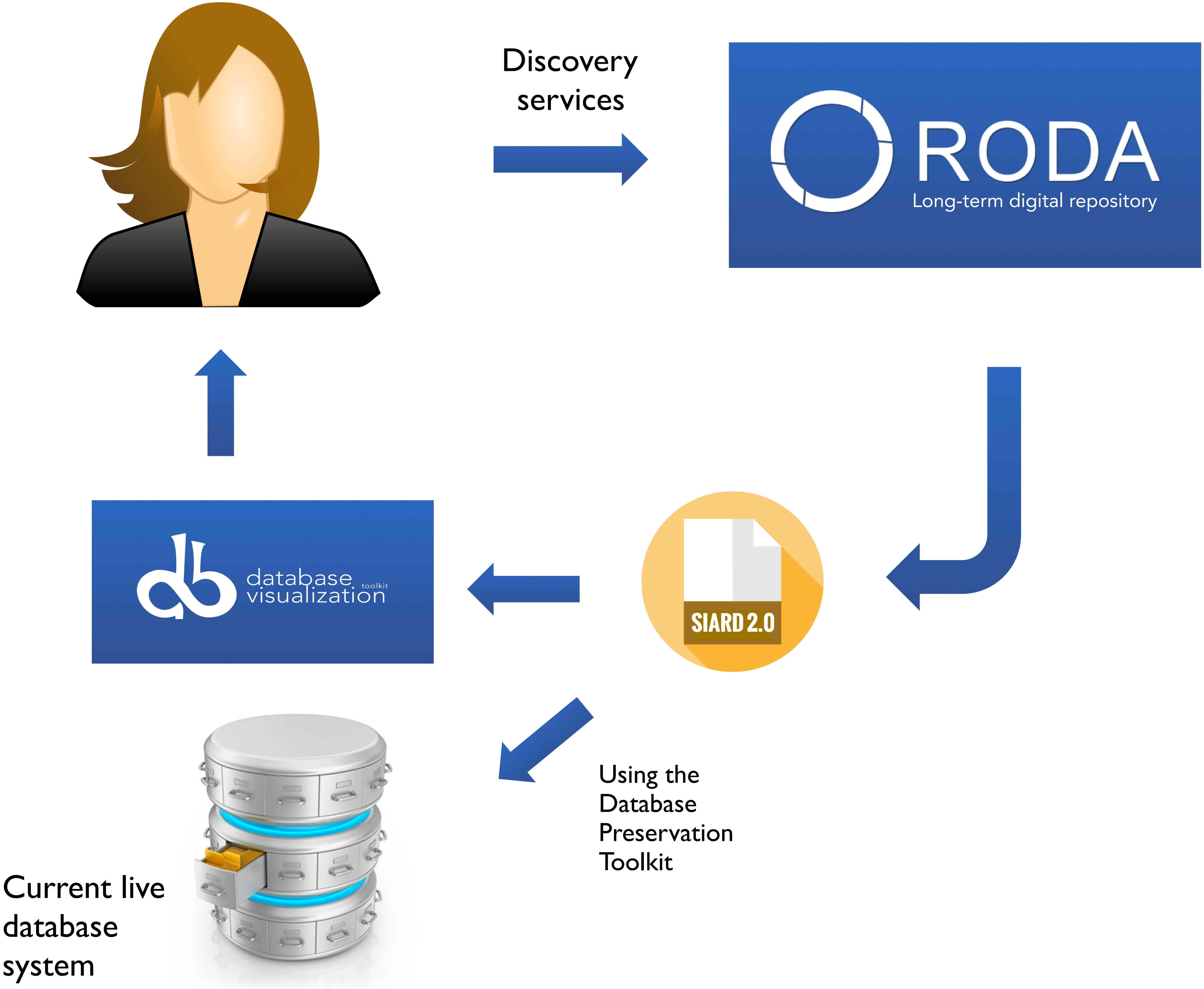


Discovery services



Discovery services

Web access to data



Discovery services

Using the Database Preservation Toolkit

Current live database system

In summary…

The Database Preservation Toolkit

Extracts data from the original RDBMS and stores it in a long-term archival format, i.e. SIARD 2

The SIARD 2 archival format

Can be submitted to a long-term preservation repository for continuous preservation, monitoring and access

Or, alternatively, can be written to a storage infrastructure

If years later, a user wants to access the archived data…

Lookup the database

Search the repository catalogue to find the right database

View database on a light-weight viewer

Using the Database Visualization Toolkit

Export parts of the database

For a shallow analysis or simple analytics using, e.g. Microsoft Excel

Load the database into a new RDBMS

Perform more advanced data analysis, e.g. OLAP

IT department is happy!

as it no longer needs to maintain the legacy database system

Management is happy!

because costs have been greatly reduced

Demo

Summary

database
preservation.toolkit

Extracts/loads data from/to relational databases

Stores data in preservation formats: SIARD 1, SIARD DK, SIARD 2

Command-line tool

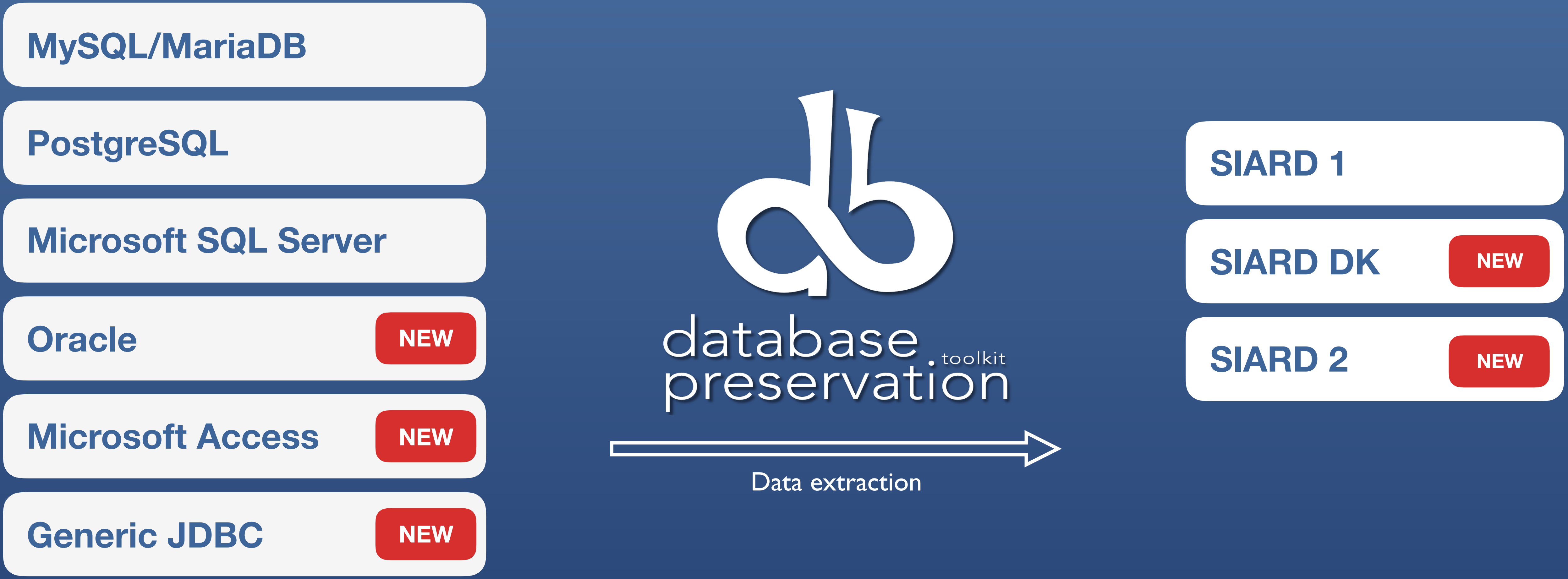
Enables using the tool over SSH

Multi-platform

Runs on Windows, Linux, MacOSX, and any Java compatible operating system

Lightweight and highly performant

Tests reveal a throughput of 88.000 records/second on a large text database



MySQL/MariaDB

PostgreSQL

Microsoft SQL Server

Oracle NEW

Microsoft Access NEW

Generic JDBC NEW

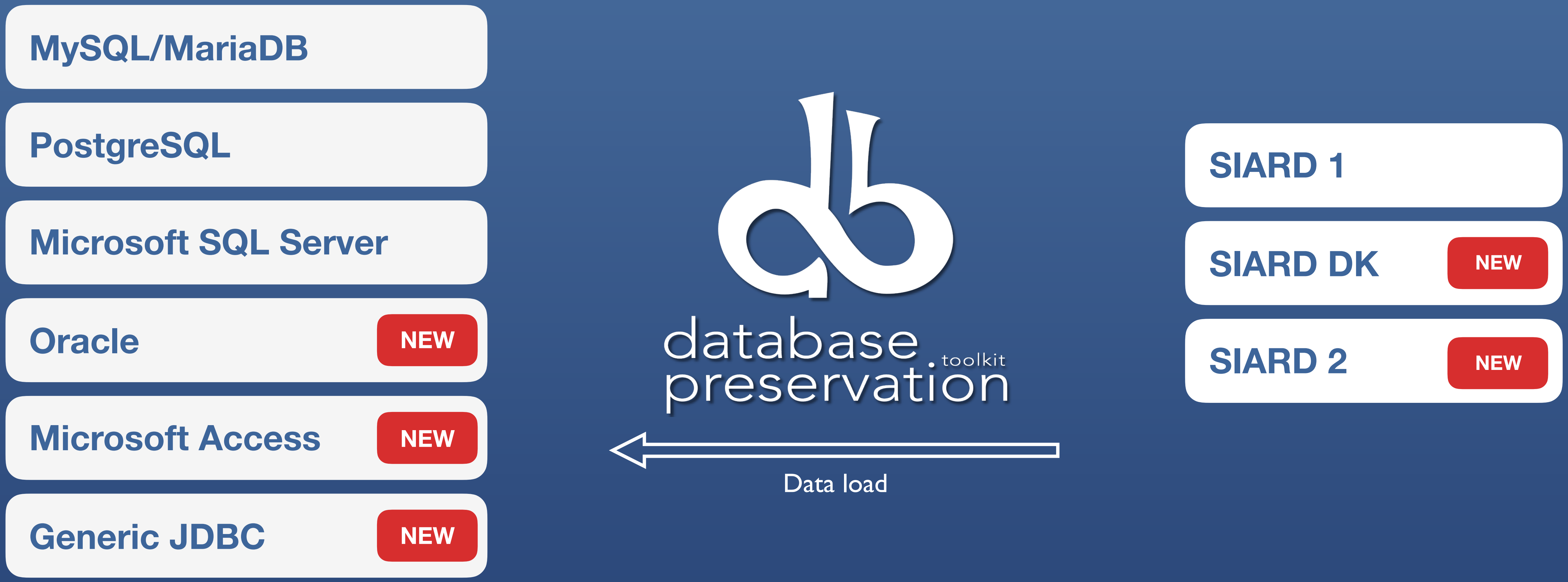
database preservation toolkit

Data extraction

SIARD 1

SIARD DK NEW

SIARD 2 NEW



MySQL/MariaDB

PostgreSQL

Microsoft SQL Server

Oracle NEW

Microsoft Access NEW

Generic JDBC NEW

database preservation toolkit

Data load

SIARD 1

SIARD DK NEW

SIARD 2 NEW

database visualization toolkit

Viewing and navigating on relational data

Lists data, follows relationships, shows structure, users, data, triggers, stored procedures, functions, metadata, etc.

Searching, filtering, column hiding, etc.

AND, OR, NOT and “exact phrase”

Export data

After filtering, the user can export data to CSV

Copes with millions of records

Supported by NoSQL horizontally scalable technologies

In conclusion…

Standards and tools are available right now

Documentation and source code published on GitHub

Examples, video tutorials are also available

Tools have been piloted in several institutions

Pilots report will be available on the E-ARK web site

http://www.eark-project.com/resources/project-deliverables/97-d25-1/file www.database-preservation.com

Questions?