Adaptor ver. 2.0

One click data export

Pietro Palladino – University of Perugia – 29th November 2008 – Brussels

Outline

- What the old adaptor did and what did not
- A new design
- The graphic configuration wizard
- A simplified export process

Adaptor general design/features

- The adaptor consists of the following parts
  1. Database connection
     - read data from a source
  2. Configuration
     - configure what to export and how to export
  3. Data exporter
     - stream data from database to the xml file

The old Adaptor design

- The adaptor consists of the following parts
  1. Database connection
     - standard Java DataBase Connectivity
  2. Simple plain configuration file
     - contains static field values
     - contains fields mapping on names
  3. Data exporter
     - database streaming
     - xml writer

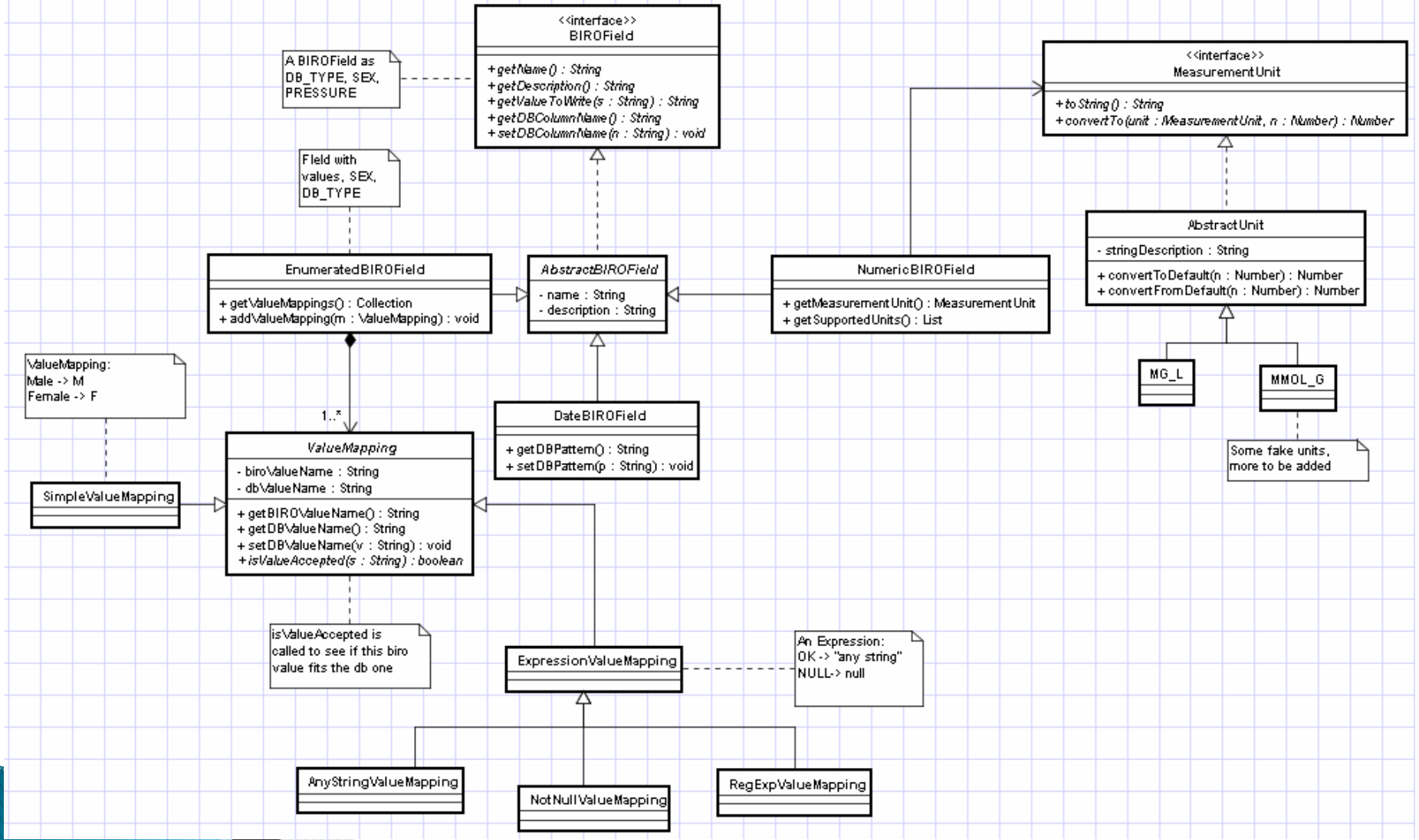
315 lines of code in 3 Java files (libraries excluded)

The Adaptor2 design

- The Adaptor2 now consists of the following parts
  1. Database connection
     - standard Java DataBase Connectivity
  2. Field mapping *(new!)*
     - applying data structure Strategy pattern
     - a class for each field capable of
       - parsing database value
       - converting to the right output value
     - high configurable and extendible (more fields can be added)
  3. Data exporter
     - database streaming
     - xml writer

2085 code lines in 90 Java files (libraries excluded)

Adaptor2 design – class diagram

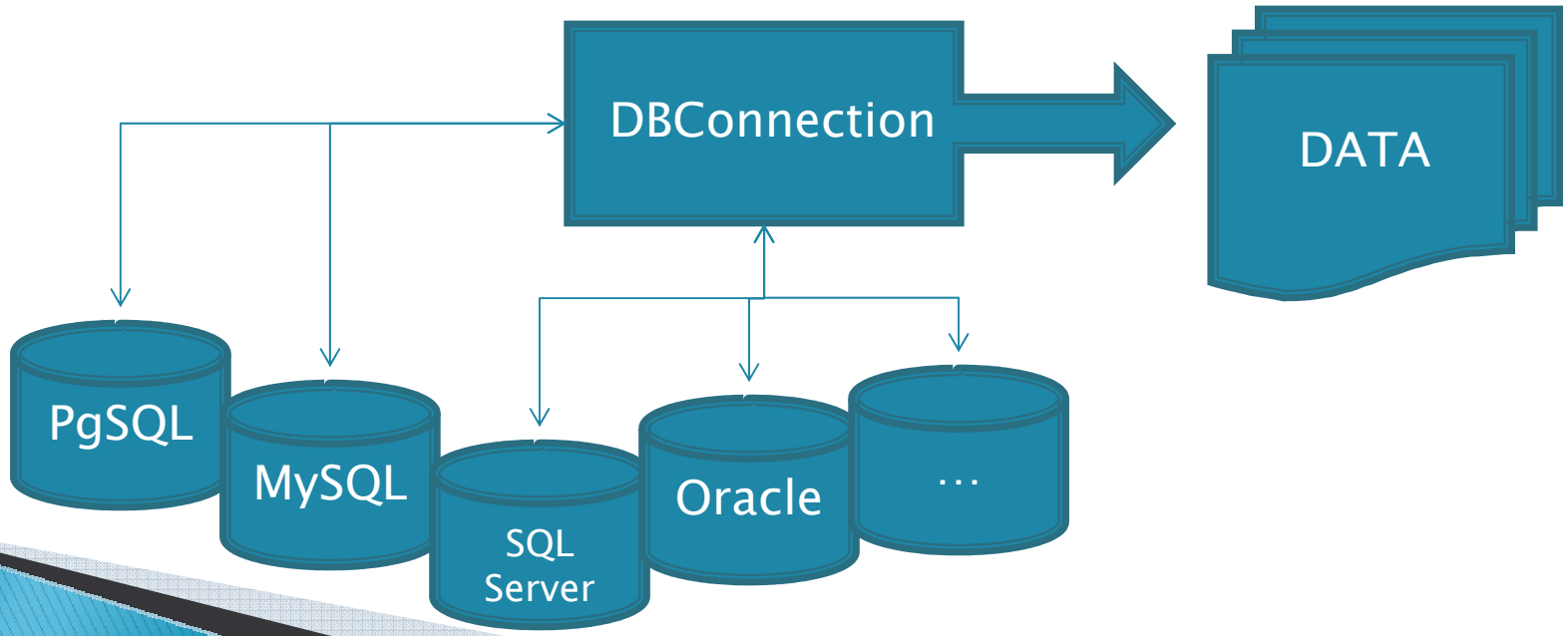


The Adaptor2 – configuration

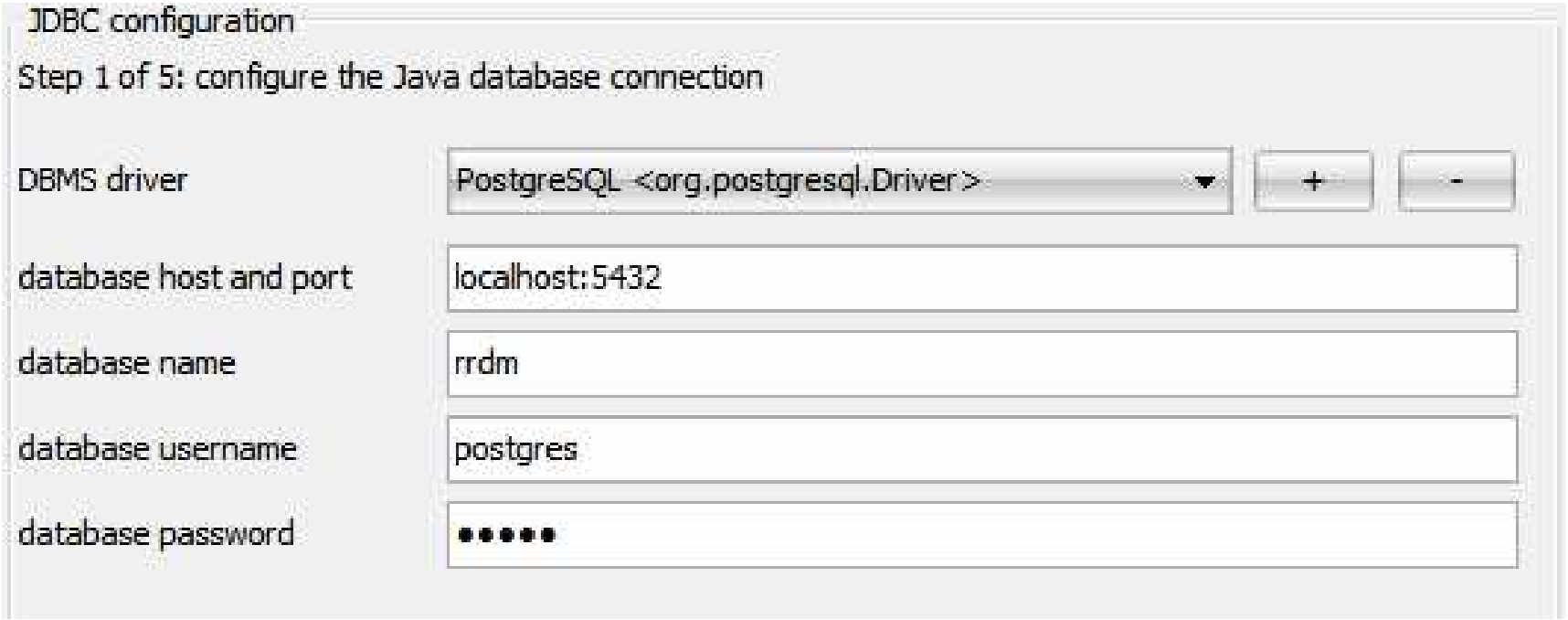
- To configure the Adaptor a set of wizards has been implemented and included in the BIRO Box GUI
  - no plain configuration files, no “format knowledge”
  - let you choose the JDBC configuration
  - let you easily map BIRO fields with database by simply selecting them from a list
  - a click on a button to start the export

The Adaptor2 – JDBC

- Database connection
  - same as old one, using JDBC
  - let you connect to many DBMSs in a common way
    - among supported: PostgreSQL, MySQL, SQLServer, Oracle, …

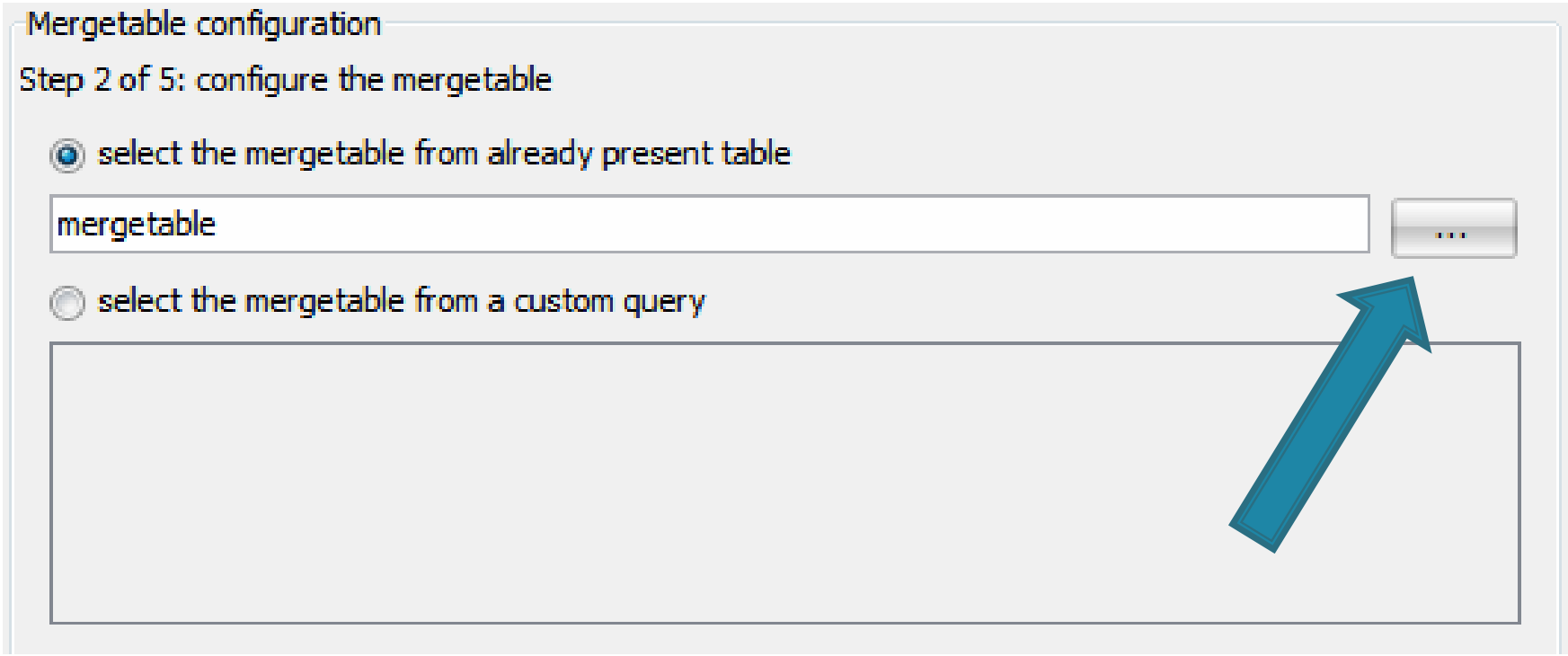


Adaptor2 wizard (step 1)



- Let you specify DBMS parameters

Adaptor2 wizard (step 2)

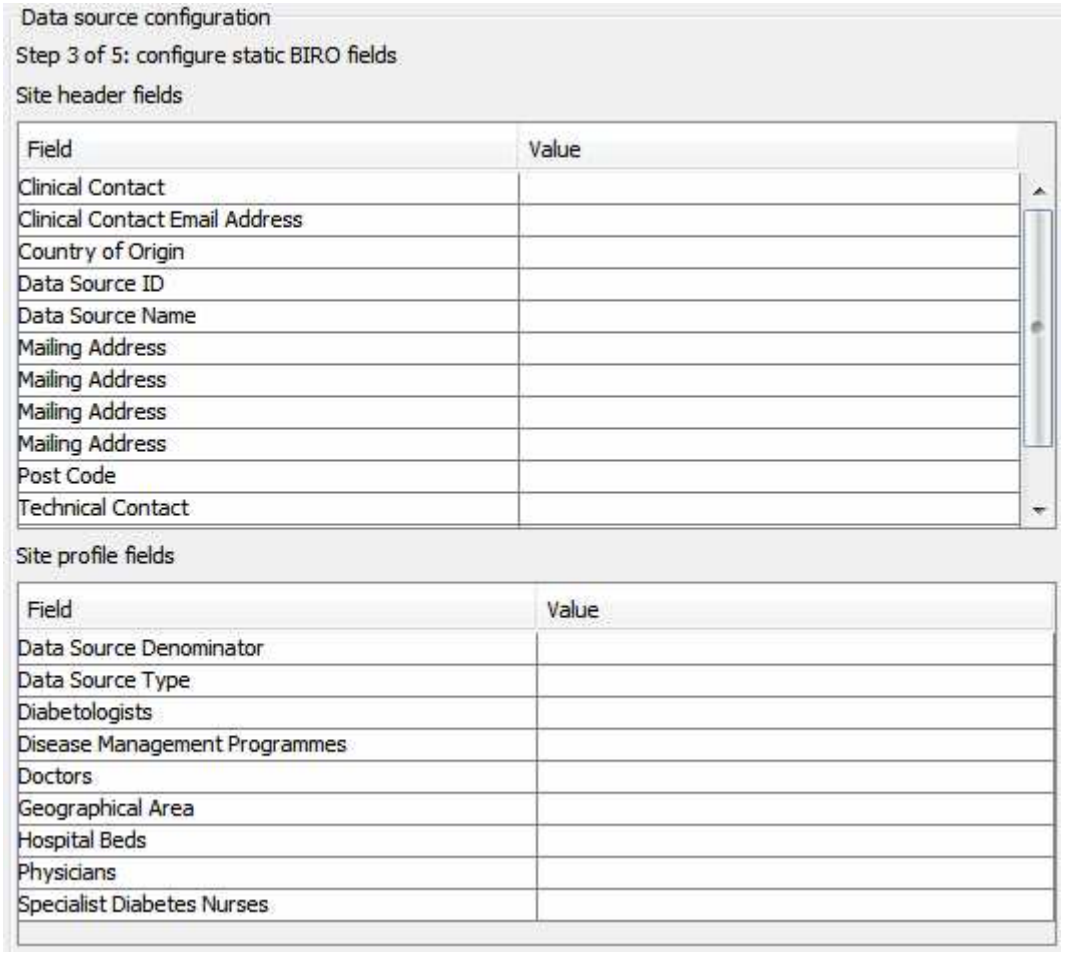


- Let you specify the table where data are stored into
- Can retrieve info directly from DB!

The Adaptor2 – Fields mapping

- Fields mapping
  - strategy pattern applied to fields makes it possible to choose which fields to export
  - developers can add more fields also at runtime
  - strategy pattern applied to numeric fields allows automatic conversion between source and target measurement units

Adaptor2 wizard (step 3)



Data source configuration

Step 3 of 5: configure static BIRO fields

Site header fields

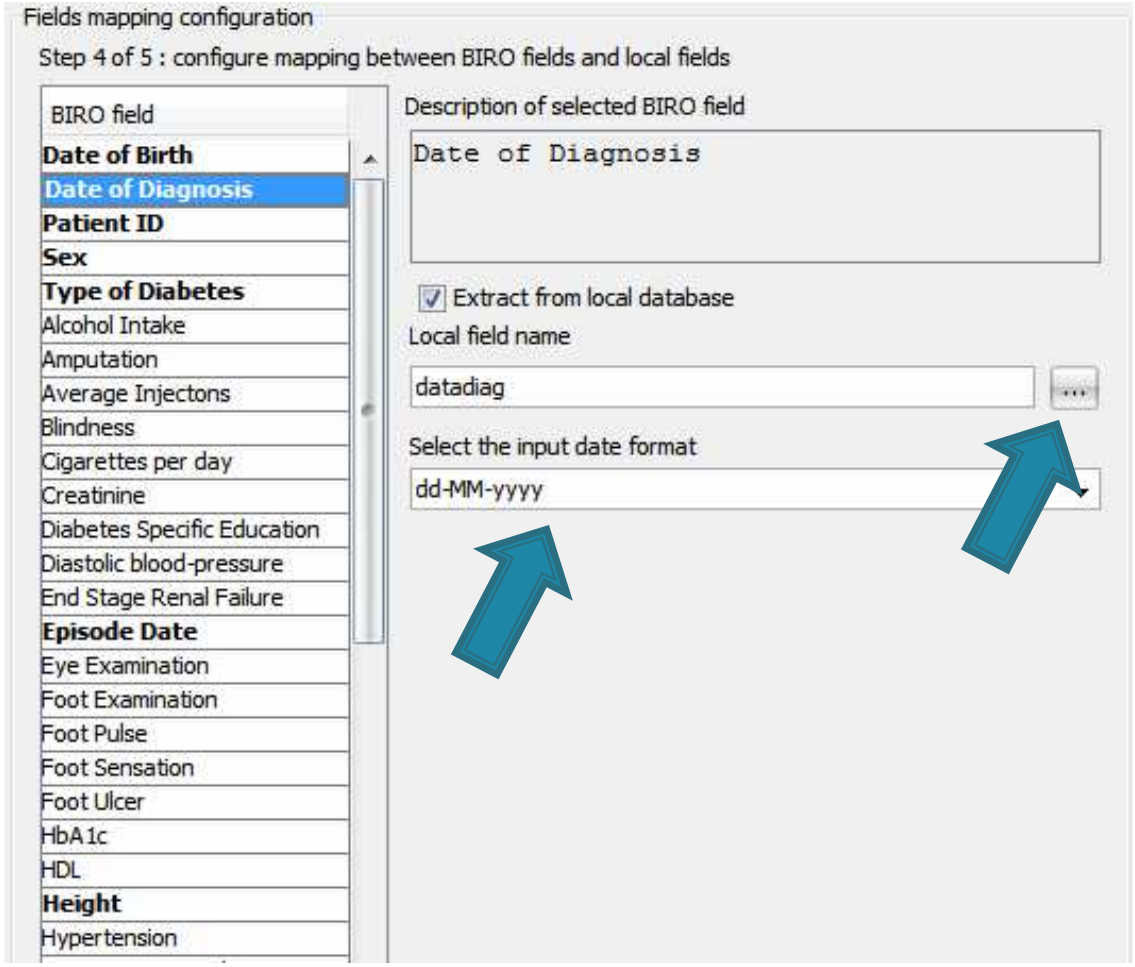
Field	Value
Clinical Contact	
Clinical Contact Email Address	
Country of Origin	
Data Source ID	
Data Source Name	
Mailing Address	
Mailing Address	
Mailing Address	
Mailing Address	
Post Code	
Technical Contact	

Site profile fields

Field	Value
Data Source Denominator	
Data Source Type	
Diabetologists	
Disease Management Programmes	
Doctors	
Geographical Area	
Hospital Beds	
Physicians	
Specialist Diabetes Nurses	

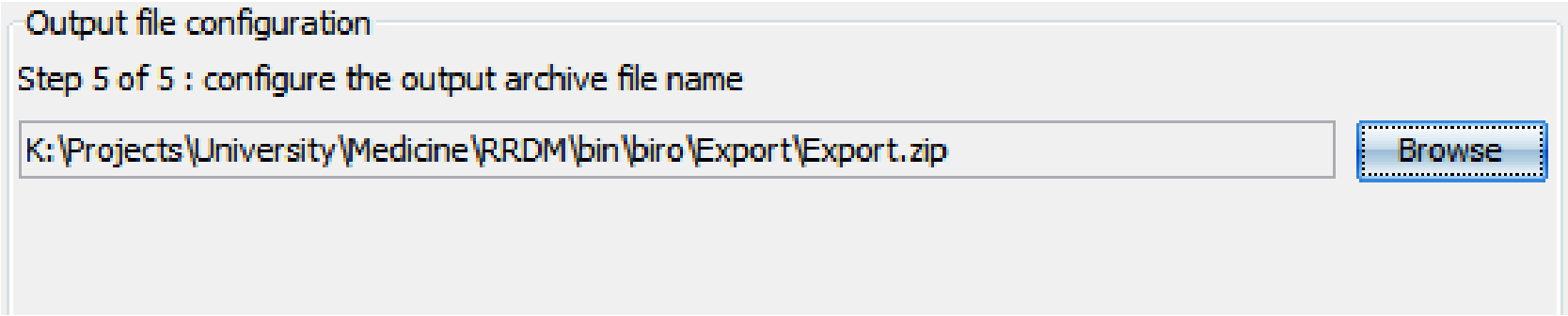
- Let you specify static data
  - data source parameters
  - next: allow these data to be read from CSV/ database

Adaptor2 wizard (step 4)



- Let you specify mapping between BIRO and local fields
  - select the fields to export
  - find field name from database
  - select options (date format, measurement unit, enumerated values and so on…)

Adaptor2 wizard (step 4)



- Let you specify the final zip file that will contain all exported xml files
- → will be sent to the BIRO Database

Run the Adaptor

- The export process takes some seconds to complete
- A progress bar and a “remaining time” label helps following the process running
- The exported archive contains (for Umbria database) over 76540 xml files
- The compressed file is less than 40MB, easier to transfer than separate files

Exported sample file (1)

```
<?xml version="1.0" encoding="Cp1252"?>
<!-- BIRO Export File -->
<ECDataExport xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance">
 <Patient>
  <Profile>
   <ProfileFieldName>PAT_ID</ProfileFieldName>
   <ProfileFieldValue>4D3882281A265A1E7496A4F59EF8704A</ProfileFieldValue>
  </Profile>
  <Profile>
   <ProfileFieldName>TYPE_DM</ProfileFieldName>
   <ProfileFieldValue>2</ProfileFieldValue>
  </Profile>
  <Profile>
   <ProfileFieldName>SEX</ProfileFieldName>
   <ProfileFieldValue>1</ProfileFieldValue>
  </Profile>
  <Profile>
   <ProfileFieldName>DOB</ProfileFieldName>
   <ProfileFieldValue>06-04-1933</ProfileFieldValue>
  </Profile>
...
```

Exported sample file (2)

```
…
  <Profile>
    <ProfileFieldName>DT_DIAG</ProfileFieldName>
    <ProfileFieldValue>01-01-1980</ProfileFieldValue>
  </Profile>
  <EpisodeData>
    <EpisodeDate>15-02-2001</EpisodeDate>
    <Data>
      <EpisodeFieldName>EPI_DATE</EpisodeFieldName>
      <EpisodeFieldValue>15-02-2001</EpisodeFieldValue>
    </Data>
    <Data>
      <EpisodeFieldName>WEIGHT</EpisodeFieldName>
      <EpisodeFieldValue>72.0</EpisodeFieldValue>
    </Data>
    <Data>
      <EpisodeFieldName>HEIGHT</EpisodeFieldName>
      <EpisodeFieldValue>1.58</EpisodeFieldValue>
    </Data>
    <Data>
      <EpisodeFieldName>SBP</EpisodeFieldName>
      <EpisodeFieldValue>170.0</EpisodeFieldValue>
    </Data>
…
```

Exported sample file (3)

```
…
      <Data>
        <EpisodeFieldName>DBP</EpisodeFieldName>
        <EpisodeFieldValue>90.0</EpisodeFieldValue>
      </Data>
      <!-- BIRO field 'HBA1C' has no value in Database -->
      <Data>
        <EpisodeFieldName>CREAT</EpisodeFieldName>
        <EpisodeFieldValue>155.0</EpisodeFieldValue>
      </Data>
      <!-- BIRO field 'MA_TEST' has no value in Database -->
      <Data>
        <EpisodeFieldName>CHOL</EpisodeFieldName>
        <EpisodeFieldValue>227.0</EpisodeFieldValue>
      </Data>
      <Data>
        <EpisodeFieldName>HDL</EpisodeFieldName>
        <EpisodeFieldValue>54.0</EpisodeFieldValue>
      </Data>
      <!-- BIRO field 'TG' has no value in Database -->
    </EpisodeData>
  </Patient>
</ECDataExport>
```

Future work

- Add JDBC auto configuration for the most common databases
- On field mappings one can think of an heuristic to automatically map BIRO fields to local fields
  - by name, by type, by value, …