



INSTITUTO NACIONAL DE ESTATÍSTICA
STATISTICS PORTUGAL



FIFTH INTERNATIONAL WORKSHOP ON BUSINESS DATA COLLECTION METHODOLOGY

19-21 SEPTEMBER 2018 – STATISTICS PORTUGAL, LISBON

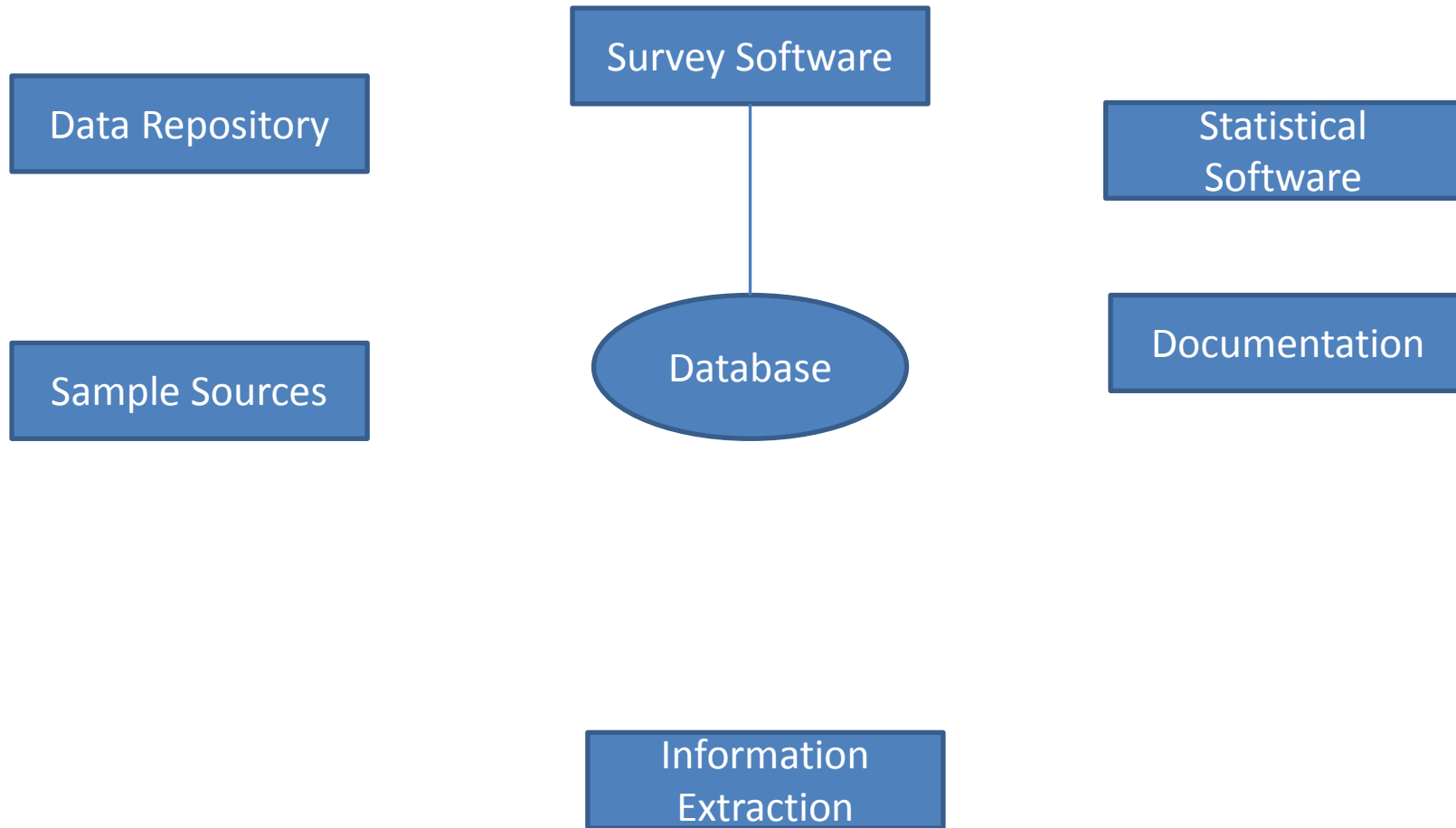




Integrating survey design and data quality management

Daniel Scheuregger, Cedefop

FIFTH INTERNATIONAL WORKSHOP ON BUSINESS DATA COLLECTION METHODOLOGY





Survey Software - Limesurvey

Questions

Conditions

Translation

Questionnaire
development

Survey Data
(Modes)

Para-Data
Collection

Participants
Management

Fieldwork
administration &
data collection

Data
Archive

Data
Exchange

Reporting

Data analysis &
archiving



Settings **Structure**

+ Add question group + Add question

- Example Group 01 (2)
 - [COUNTRY] › Country of residence
 - [QExample02] › Using this card, which of th...
- Example Group 02 (1)
 - [QExample03] › I am going to read out a li...

Preview survey Preview question group Preview question Edit Check logic

Delete Export Copy Set conditions Edit subquestions Edit answer options

Question summary *QExample03* (ID: 657)

Question group: *Example Group 02* (ID:28)
Code: *QExample03 : (Optional question)*
Question: I am going to read out a list of things about your working life. Using this card, please say ... READ OUT...
Help:
Type: Array
Mandatory: No
Relevance equation: 1

Question quick actions

Add new question to group

+

Add new question to group



3 QExample03 I am going to read out a list of things about your working life. Using this card, please say how much the management at your work allows/allowed you...


READ OUT...

	no influ- ence	2	3	4	5	6	com- plete control	No an- swer
...to decide how your own daily work is/was organised?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input checked="" type="radio"/>
...to choose or change your pace of work?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input checked="" type="radio"/>

Example based on ESS



Translate survey

Translat
to: 

Translate survey

Survey title and description

Welcome and end text

Question groups

Questions

Subquestions

Answer options

Invitation email subject

Reminder email subject

Confirmation email subject

Registration email subject

Question code / ID	English	Portuguese
COUNTRY (651)	Country of residence	<div style="border: 1px solid gray; padding: 5px;"> <p><u>País de residência</u></p> </div>
QExample02 (652)	Using this card, which of these descriptions applies to what you have been doing for the last 7 days? Select all that apply.	<div style="border: 1px solid gray; padding: 5px;"> <p>Translation</p> </div>

FIFTH INTERNATIONAL WORKSHOP ON BUSINESS DATA COLLECTION METHODOLOGY



File Edit Search Tools Go to Help

Database filter Table filter Host: 127.0.0.1 Database: test Table: questions Data Query*

test.questions: 131 rows total

qid	parent_qid	sid	gid	type	title	question
367	0	296,897	11	M	FEDCHAR	<p> Please indicate
359	0	296,897	11	L	FEDFIELD	<p> What was the fi
169	0	296,897	9	N	REGION	Region of Residence (N
408	0	296,897	10	Y	UNEMP01	<p> Have you b
356	0	296,897	11	L	FEDLEVEL	<p> What was tl
357	0	296,897	11	L	FEDCOMP	<p> Have you comp
346	0	296,897	10	Y	GUIDEINSTA	<p> During the
170	0	296,897	9	L	DEGURB	Degree of urbanisation
434	0	296,897	22	Y	TRNEED	Overall, would you say
448	0	296,897	23	Y	WRKENV04	<span style="font-size:
449	0	296,897	23	Y	WRKENV05	<span style="font-size:
450	0	296,897	23	Y	WRKENV06	<span style="font-size:
451	0	296,897	23	Y	WRKENV07	<span style="font-size:
364	171	296,897	9	K	SQ003	other
362	171	296,897	9	K	SQ001	Adults
363	171	296,897	9	K	SQ002	Children less than 25 y
352	348	296,897	10	T	SQ004	Other type of informati
349	348	296,897	10	T	SQ001	Information or advice/f
350	348	296,897	10	T	SQ002	Assessment of skills an
351	348	296,897	10	T	SQ003	Information or advice/f
353	348	296,897	10	T	SQ005	No Answer
368	367	296,897	11	T	SQ001	Was mainly distance le
369	367	296,897	11	T	SQ002	Included in person teac
370	367	296,897	11	T	SQ003	was organised in terms
371	367	296,897	11	T	SQ004	was a regular lecture st
372	367	296,897	11	T	SQ005	I had contact to other p
373	367	296,897	11	T	SQ006	I had contact to other p
417	374	296,897	14	O	SQ001	Courses or workshops

Database filter Table filter

- old_tokens_483942_20160/04160011 3.2 KiB
- old_tokens_483942_20160706133838 3.3 KiB
- old_tokens_483942_20170419142203 7.1 KiB
- participants 2.2 KiB
- participant_attribute 1.0 KiB
- participant_attribute_names 2.0 KiB
- participant_attribute_names_lang 2.0 KiB
- participant_attribute_values 1.0 KiB
- participant_shares 1.0 KiB
- permissions 9.9 KiB
- plugins 2.3 KiB
- plugin_settings 1.0 KiB
- questions 35.2 KiB
- question_attributes 4.4 KiB
- quickmenu_sortorder 2.0 KiB
- quota 1.0 KiB
- quota_languagesettings 1.0 KiB
- quota_members 1.0 KiB
- saved_control 1.0 KiB
- sessions 1.0 KiB
- settings_global 9.0 KiB
- settings_user 2.1 KiB
- surveymenu 11.2 KiB
- surveymenu_entries 20.8 KiB
- surveys 2.4 KiB
- surveys_groups 2.1 KiB
- surveys_languagesettings 12.7 KiB
- survey_links 1.0 KiB
- survey_url_parameters 1.0 KiB
- templates 17.0 KiB
- template_configuration 15.8 KiB
- tutorials 4.0 KiB
- tutorial_entries 1.0 KiB
- tutorial_entry_relation 1.0 KiB
- users 5.7 KiB
- user_groups 3.0 KiB

FIFTH INTERNATIONAL WORKSHOP ON BUSINESS DATA COLLECTION METHODOLOGY



Host: 127.0.0.1 Database: test Table: questions Data Query*

```

1 select qid, parent_qid, sid, title, question from test.questions
2 where parent_qid = '171' XOR qid= '171';
3
4
    
```

Variable ID Survey Variable name Question text

qid	parent_qid	sid	title	question
171	0	296,897	HHNBERS	<p> </p><p>How many people live ...
364	171	296,897	SQ003	other
362	171	296,897	SQ001	Adults
363	171	296,897	SQ002	Children less than 25 years

Host: 127.0.0.1 Database: test Table: questions Data Query*

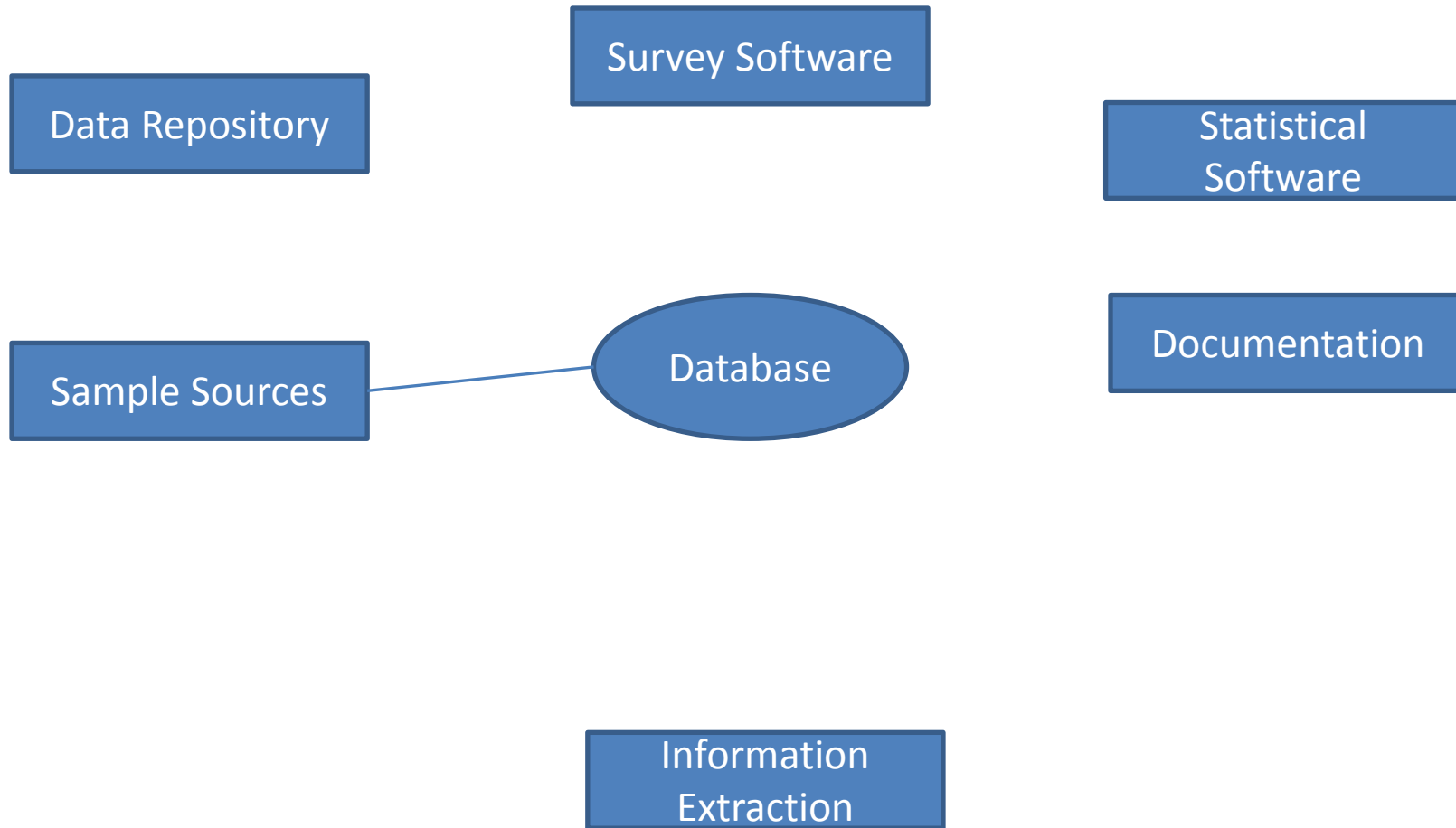
```

1 select qid, parent_qid, sid, title, question, tag01 from test.questions
2 where tag01 Like "socio%";
3
4
    
```

Added information

qid	parent_qid	sid	title	question	tag01
343	0	296,897	BIRTHMOTHER	What is your mother's country of bi...	sociodemographic
175	0	296,897	CITIZEN	What is your citizenship? <s...	sociodemographic
179	0	296,897	BIRTHPLACE	What is your country of birth?<br /...	sociodemographic
181	0	296,897	MARSTADEFACTO	What is your de facto marital status...	sociodemographic
295	0	296,897	MAINSTAT	Do you carry out a job or professio...	sociodemographic

FIFTH INTERNATIONAL WORKSHOP ON BUSINESS DATA COLLECTION METHODOLOGY



FIFTH INTERNATIONAL WORKSHOP ON BUSINESS DATA COLLECTION METHODOLOGY



Central participant database

[Display CPDB participants](#) [Info](#) [Import](#) [Blacklist settings](#) [Attributes](#) [Share panel](#) [Export all participants](#)

Selection criteria

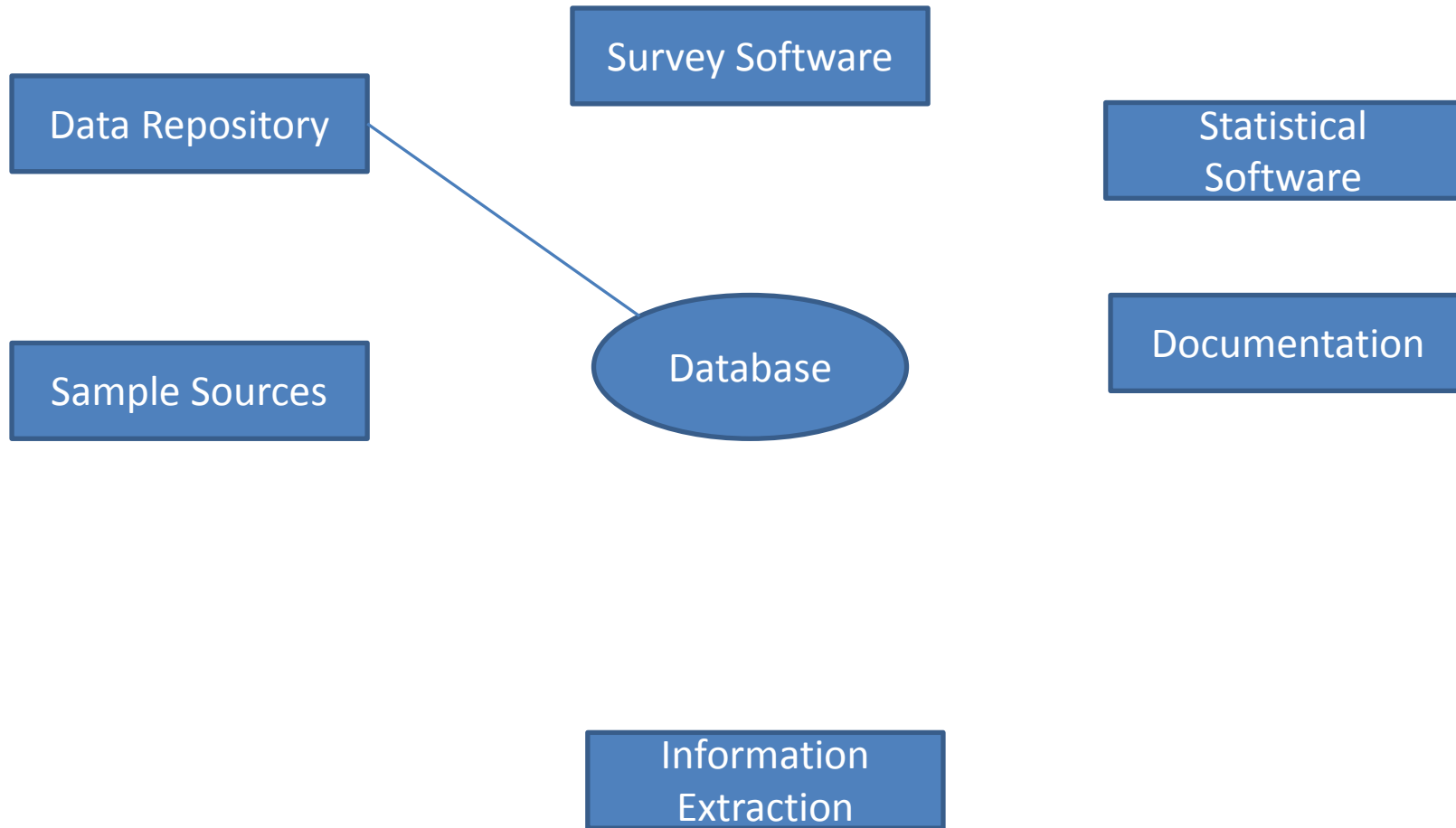
<input type="checkbox"/> Action	Last name <input type="text"/>	First name <input type="text"/>	Email address <input type="text"/>	Language <input type="text"/>	Active surveys <input type="text"/>	Owner <input type="text"/>
<input type="checkbox"/>	Scheuregger	Daniel	daniel.scheuregger@cedefop.europa.eu			

Participants list

Added criteria

Blacklisted <input type="text"/>	Created on <input type="text"/>	country <input type="text"/>	New Attribute <input type="text"/>	New Attribute <input type="text"/>
<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>
<input type="checkbox"/> No	13.08.2018		attribute entry	another attribute

FIFTH INTERNATIONAL WORKSHOP ON BUSINESS DATA COLLECTION METHODOLOGY





test

old_survey_125612_20180813144902

startlanguage	(...)	125612X29X651	125612X29X652SQ001	125612X29X652SQ006	125612X29X652SQ005
en		00	1	0	0

Variables

Responses

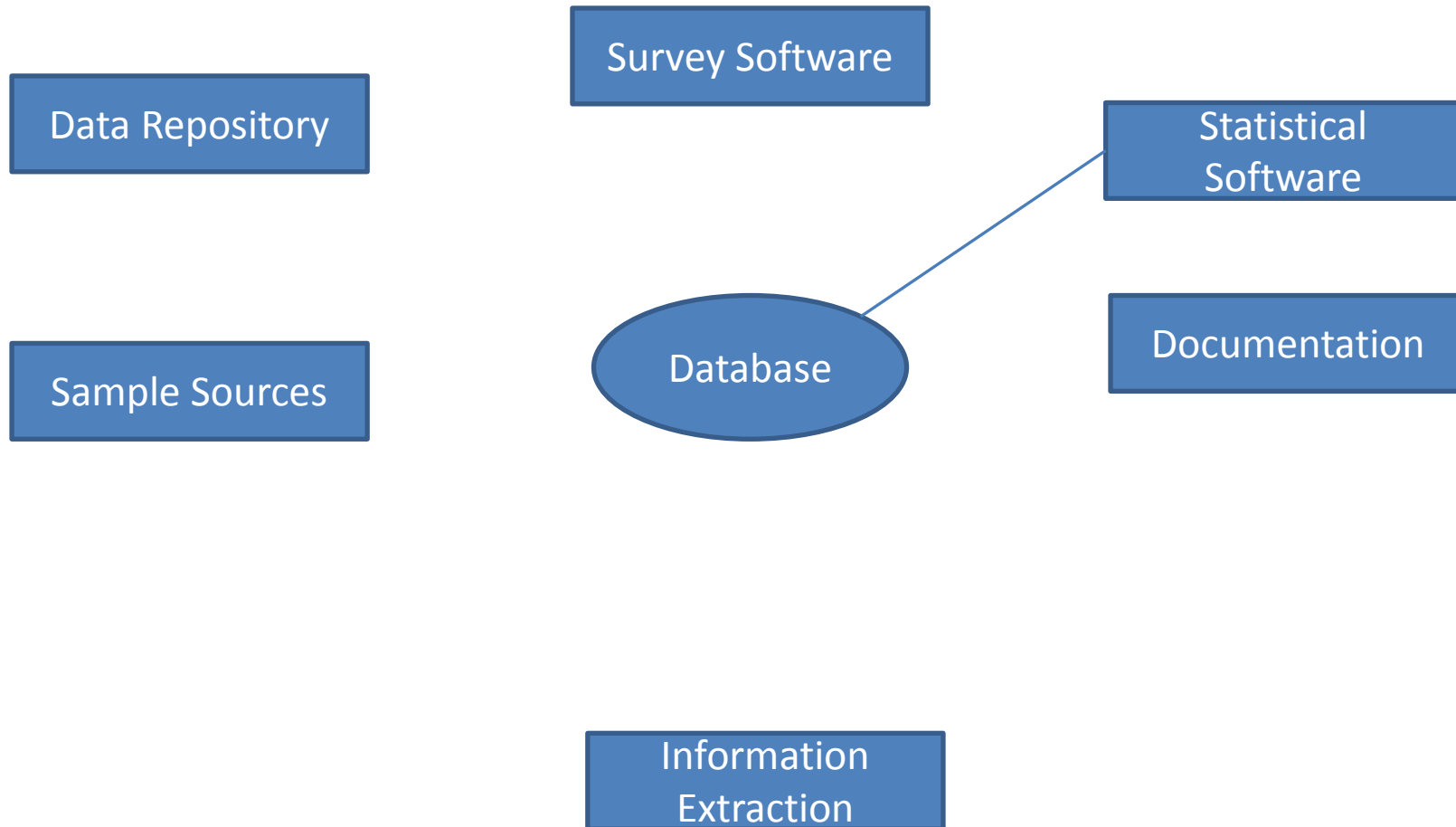
[
.
.
.
]

Survey Group Question

125612X29X651

Delimiter

FIFTH INTERNATIONAL WORKSHOP ON BUSINESS DATA COLLECTION METHODOLOGY



R Console

Platform: x86_64-w64-mingw32/x64 (64-bit)

R is free software and comes with ABSOLUTELY NO WARRANTY. You are welcome to redistribute it under certain conditions. Type 'license()' or 'licence()' for distribution details.

Natural language support but running in an English locale.

R is a collaborative project with many contributors. Type 'contributors()' for more information and 'citation()' on how to cite R or R packages in publications.

Type 'demo()' for some demos, 'help()' for on-line help, or 'help.start()' for an HTML browser interface to help. Type 'q()' to quit R.

```
> source("C:\\Users\\...\\Desktop\\test.R")
```

Warning messages:

1: package 'RMariaDB' was built under R version 3.4.4

2: package 'DBI' was built under R version 3.4.4

3: In result_create(conn@ptr, statement, is_statement) :
Cancelling previous query

```
> View(d1_questions)
```

```
> View(d2_Responses)
```

```
> |
```

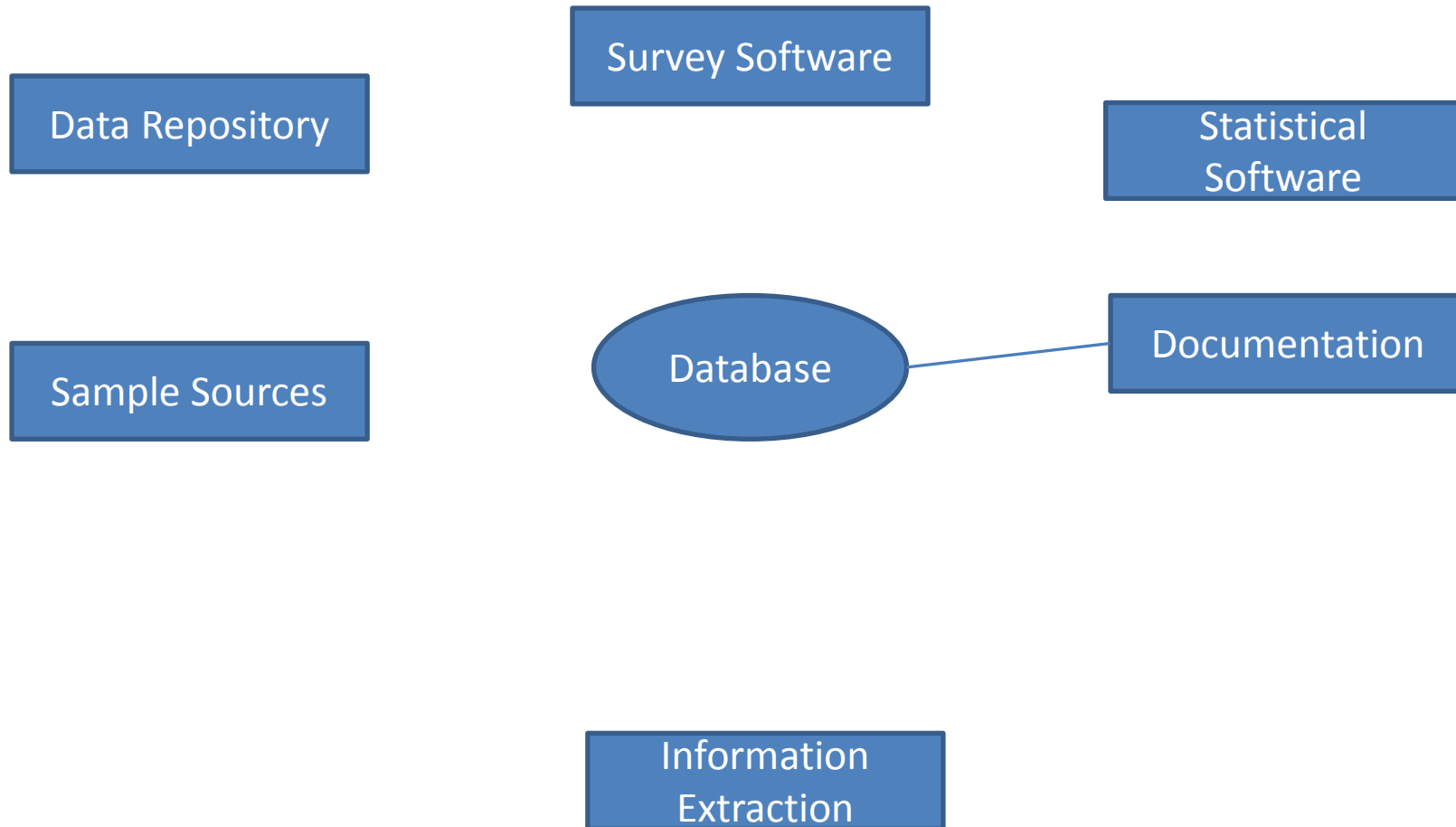
R Data: d1_questions

	qid	parent_qid	sid	gid	type	title
1	598	597	296897	11	T	SQ003
2	599	597	296897	11	T	SQ002
3	609	608	296897	11	T	SQ003
4	173	0	296897	9	L	SEX
5	174	0	296897	9	D	Birthdate
6	343	0	296897	9	!	BIRTHMOTHER
7	175	0	296897	9	!	CITIZEN
8	179	0	296897	9	!	BIRTHPLACE
9	180	0	296897	9	N	RESTIME
10	181	0	296897	9	L	MARSTADEFACTO

R Data: d2_Responses

	startlanguage	125612X29X651	125612X29X652SQ001	125612X29X652SQ006
1	en	00	1	0

FIFTH INTERNATIONAL WORKSHOP ON BUSINESS DATA COLLECTION METHODOLOGY



FIFTH INTERNATIONAL WORKSHOP ON BUSINESS DATA COLLECTION METHODOLOGY



Host: 127.0.0.1 Database: test Table: documentation_dataproc... Data Query

test.documentation_dataprocessing: 1 rows total (approximately) Next Show all Sorting (1)

sid	RSyntax	StataSyntax
125612		(NULL)

Text editor

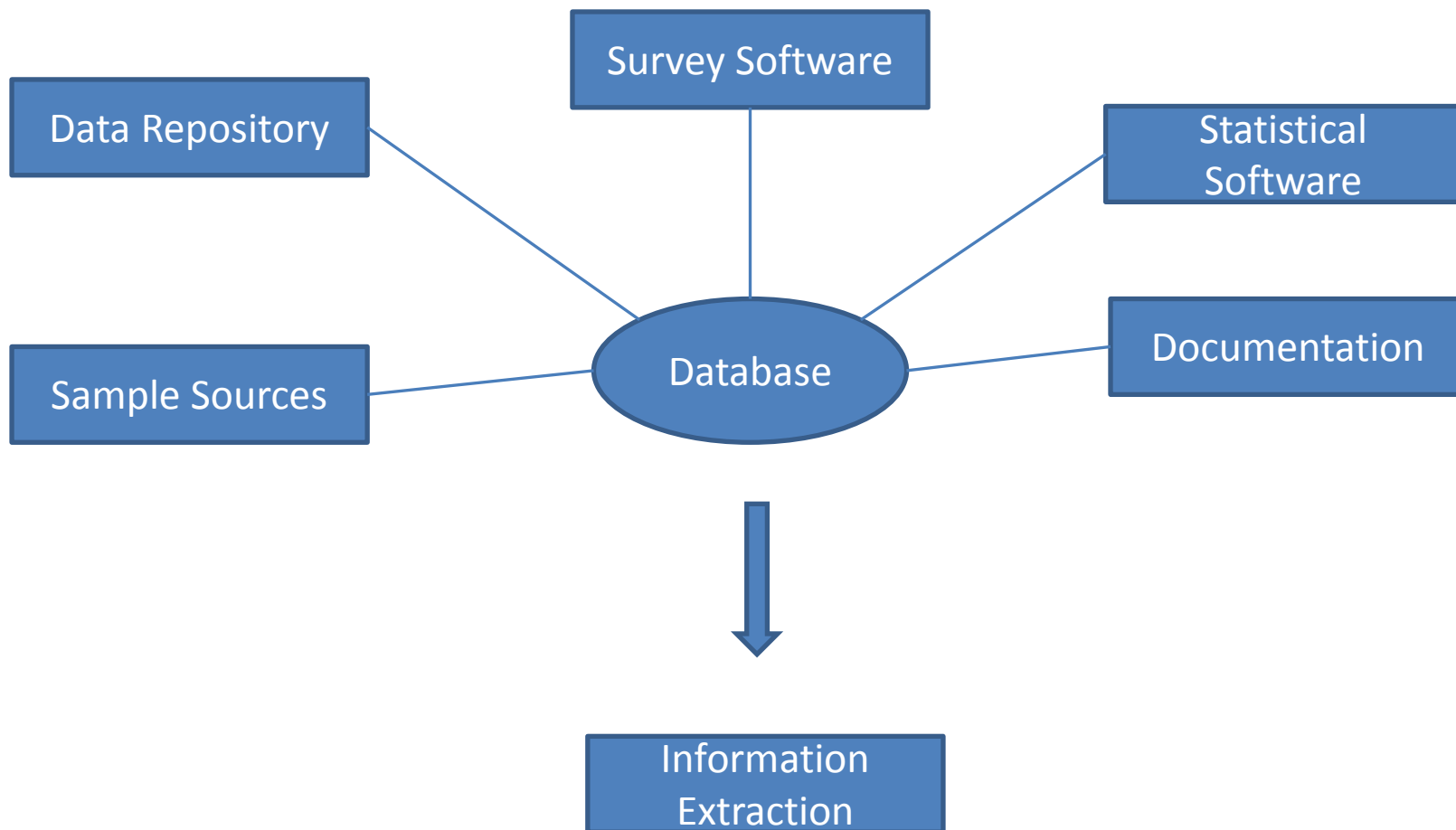
```
library(RMariaDB)
library(DBI)

con <- dbConnect(RMariaDB::MariaDB(), dbname = "test",
username = "root", password = "***", host = NULL,
bigint = c("integer64", "integer", "numeric", "character") )

# query
rs2 <- dbSendQuery(con, "SELECT startlanguage, 125612X29X651, 125612X29X652SQ001, 125612X29X652SQ006
FROM test.old_survey_125612_20180813144902")

d2_Responses <- dbFetch(rs2, n = 10) # extract data in chunks of 10 rows
```

462 characters.





Thank you

Daniel.SCHEUREGGER@cedefop.europa.eu

FIFTH INTERNATIONAL WORKSHOP ON BUSINESS DATA COLLECTION METHODOLOGY

