



Datamanagement

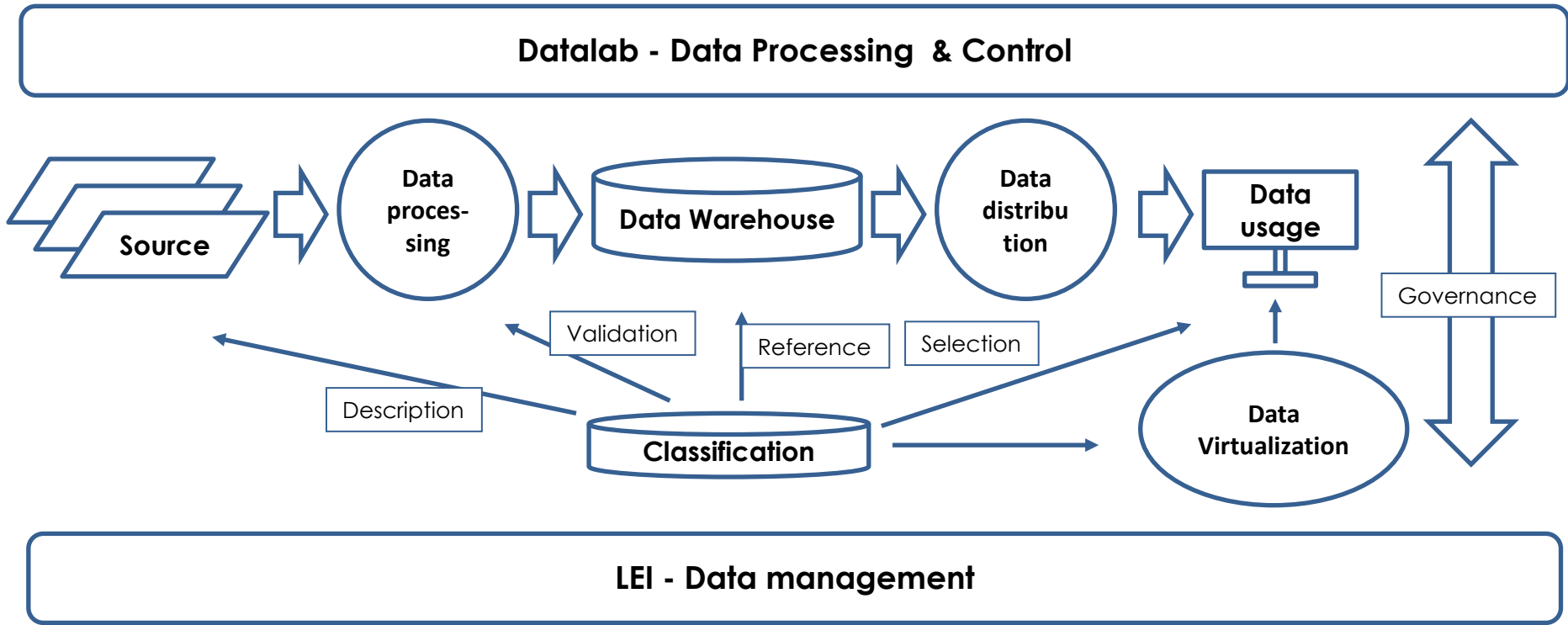
With applications in
FLINT (partly)
and FADN (current and future)

Pacioli, Sept. 28

Pristina, Kosovo

Derek Verweij & Eugene Westerhof

Overview

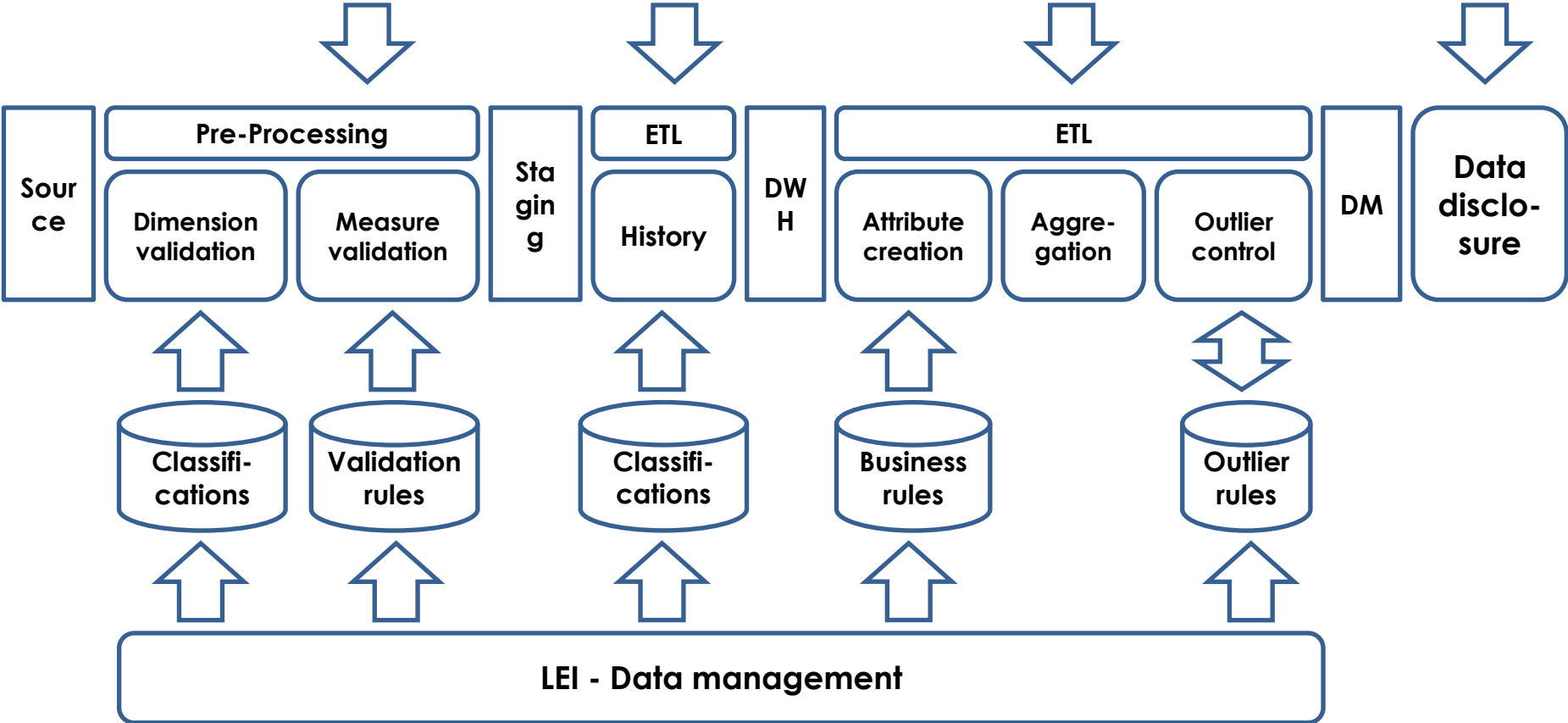


Project Scope

Datawarehousing

- Congebrate data from multiple sources into a single database so a single query engine can be used to present data.
- Integrate data from multiple source systems, enabling a central consistent view (classification)
- Maintain [data history](#), even if the source transaction systems do not (performance, reproducibility)
- Improve [data quality](#), by providing consistent codes and descriptions, flagging (or even fixing) bad data and reducing manual processing

Datalab - Data Processing & Control



Data Dimensions: Classifications

Subject:
Other Gainful Activity

Variable:
Production

Unit:
Quintals

Type:
Processing of cow's milk

103

Year:
2015

Country:
Hungary

Holding:
768.7.75195

L_PR_261_MD_Q

Current state of affairs

- First implementation of technical solution delivered
- Three pilots in three different domains:
 - Foodprofiler. Smart phone app ('gamified') to collect data on consumption patterns (daily)
 - KCB; dutch export data on agricultural products
 - BIN (Dutch FADN+ data on farm level). Data released for research updated daily
- Rule Component used for FLINT

Automatic validation feedback

The screenshot shows the Outlook interface for the mailbox 'fb.lei@wur.nl'. The ribbon includes 'Respond', 'Quick Steps', 'Move', 'Tags', and 'Find'. The left pane shows a list of emails, with the most recent one selected: 'Leij, FB RE: BINSQL validation errors ... Holding Message' at 16:34. The main pane displays the email content, which includes a table of error details.

Search Current Mailbox (Ctrl+...) Current Mailbox

Folder View McAfee E-mail Scan Tell me what you want to do...

Reply Reply All Forward Meeting IM More

Move to: ? To Manager Done Create New

Team Email Reply & Delete

Move Rules OneNote Unread/Read Categorize Follow Up

Search People Address Book Filter Email

Respond Quick Steps Move Tags Find

Search Current Mailbox (Ctrl+...) Current Mailbox

All Unread By Date Newest

Today

Leij, FB
RE: BINSQL validation errors ... 16:34
Holding Message

fb.lei@wur.nl
SSIS package loaded succes... 11:45
Please check the SSISDB report

fb.lei@wur.nl
SSIS package loaded succes... 6:38
Please check the SSISDB report

ma 26-9-2016 16:34

Leij, FB

RE: BINSQL validation errors occurred on the environment: Production


To Leij, FB

Holding	Message	Reference	Time
16372	Execution of business rule failed with error: Sequence contains no holder	HoldingStructure.OldestHolder	9/26/2016 6:13:00 AM

Farmer feedback website

https://www.3.lei.wur.nl/BIIDeelnemersSite/Resultaat.aspx - Internet Explorer
https://www.3.lei.wur.nl/BIIDeelnemersSite/Resultaat.aspx www.3.lei.wur.nl

File Edit View Favorites Tools Help

 **WAGENINGEN**
UNIVERSITY & RESEARCH

Mijn agrimatie

Home Welkom Resultaat Over ons Contact Log out Help

Mijn resultaten

- Structuur
- Bedrijfsresultaat
- Sectorresultaat
- Milieu

Bedrijfsvergelijking

- Mijn groep en mijn bedrijf
- Andere groepen
- Gewassen
- Vee

PDF versies

- Bedrijfsverslag
- Bedrijfsvergelijkend overzicht
- Saldo van gewassen
- Saldo veehouderij
- Glastuinbouwbedrijven

Overig

- Risicobarometer
- Privacy
- Handige links

Kies een tabel

- **Kostprijs melk**
- Resultaat melkvee
- Resultaat graasvee
- Beweiding
- Voeropname en productie
- BEX

Kies een jaar

- 2015
- 2014
- 2013
- 2012
- 2011
- 2010
- 2009
- 2008
- 2007
- 2006
- 2005
- 2004
- 2003

Kostprijs melk (excl. BTW) - Bedrijf 16005

De kosten (incl berekende kosten) worden eerst uitgedrukt per 100 kg melk en daarna op basis van de verdeling van de opbrengsten toegerekend naar het product melk (kostprijs). De gegevens worden alleen getoond voor zeer zuivere melkveebedrijven. Als gegevens niet worden getoond, dan is het belang van tweede tak of verbreding (inclusief zelfzuivelen) te groot om de kostprijs te kunnen bepalen.

Variabelen	Eenheid	2015	2014	2013
Totale kosten per 100 kg melk		49,62	56,77	55,16
Toegerekende kosten	euro	17,42	15,10	16,69
Arbeid en loonwerk	euro	15,07	15,18	14,79
Rente, pacht en afschrijving	euro	9,63	18,49	17,68
Overig	euro	7,49	8,00	6,00
Totale opbrengsten per 100 kg melk		37,83	42,02	47,01
Melk- en zuivelopbr. na aftrek superheffing	euro	33,44	40,86	41,72
Niet melk opbrengsten (excl. subsidies)	euro	4,39	1,16	5,29
Resultaat				
Marge (=netto-bedrijfsresultaat)	euro	-11,79	-14,75	-8,15
Kostprijs melk, per 100 kg	euro	43,86	55,21	48,95

Export naar pdf Alle gegevens

De resultaten in dit overzicht zijn afkomstig uit het Bedrijfsinformatienet van het LEI. De gegevens zijn vertrouwelijk en mogen niet worden gedeeld met derden. De bedragen zijn afzonderlijk afgerond, waardoor kleine verschillen kunnen zijn ontstaan in de totaaltellingen. Het LEI heeft dit overzicht met zorg samengesteld, maar kan niet verantwoordelijk worden gehouden voor transmissiefouten. Voor opmerkingen, wijzigingen of vragen kunt u zich wenden tot uw contactpersoon, Hendrik de Vries

Validation Rules

RICA

If L_CV_261:321_MD_V > 0
then
L_OV_261:321_MD_V > 0
or
L_PR_261:321_MD_Q > 0

FLINT

if (Type == {Processing of cow's milk}
|
{Processing of buffalo's milk}
|
{Processing of sheep's milk}
|
{Processing of goat's milk})
&
ClosingValuation > 0

then Production > 0
|
OpeningValuation > 0

FLINT

- 170 variables and 60 dimensions in 40 'subjects' (tables) (51000 variables in RICA1)
- All (~200) FLINT coherence tests implemented
- 435 RICA test automatically generated from rule configuration XML provided by the Commission
- 63 tests not generated because of missing information and different processing of lookups (= dimensions).
- Runs all tests for 100 farms is less than half a minute (RICA1 test server takes 2.5 hours)
- Rule engine Software is free for use and we welcome an open source community effort

Future developments

- Wageningen Economic Research needs to find new funding: government funding will decrease, PPS will have to pick up as will contract research
- What is our market proposition?
 - Research, plus combination of various data
- So: Agro Economic Platform in the making
- Data Warehouse will accommodate multiple datasources, including third party sources

Things to be considered

- Classification is key!
- Usability: researchers will have to easily 'construct' their information, minimal dependency on Datalab department
- Open data: is a challenge in terms of marketing (what is our market proposition?)
- Virtual data – just in time (cost, actuality)
- Big data: being developed on university level

Thank you!

- Questions