

Kettle

Или трансформация данных чайником

Igor Bossenko

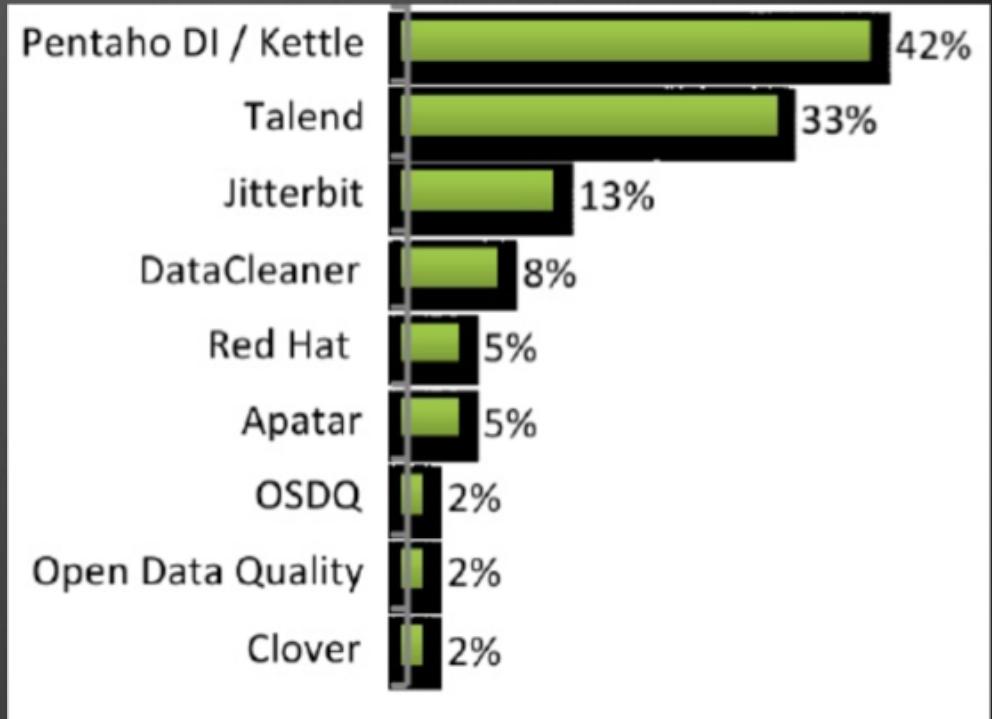
32.06.2015

Why Data Integration Tool?

- Data consolidation project for Nortal BI
 - 6 countries
 - 11 different applications
 - 3 DBMS

Why Kettle?

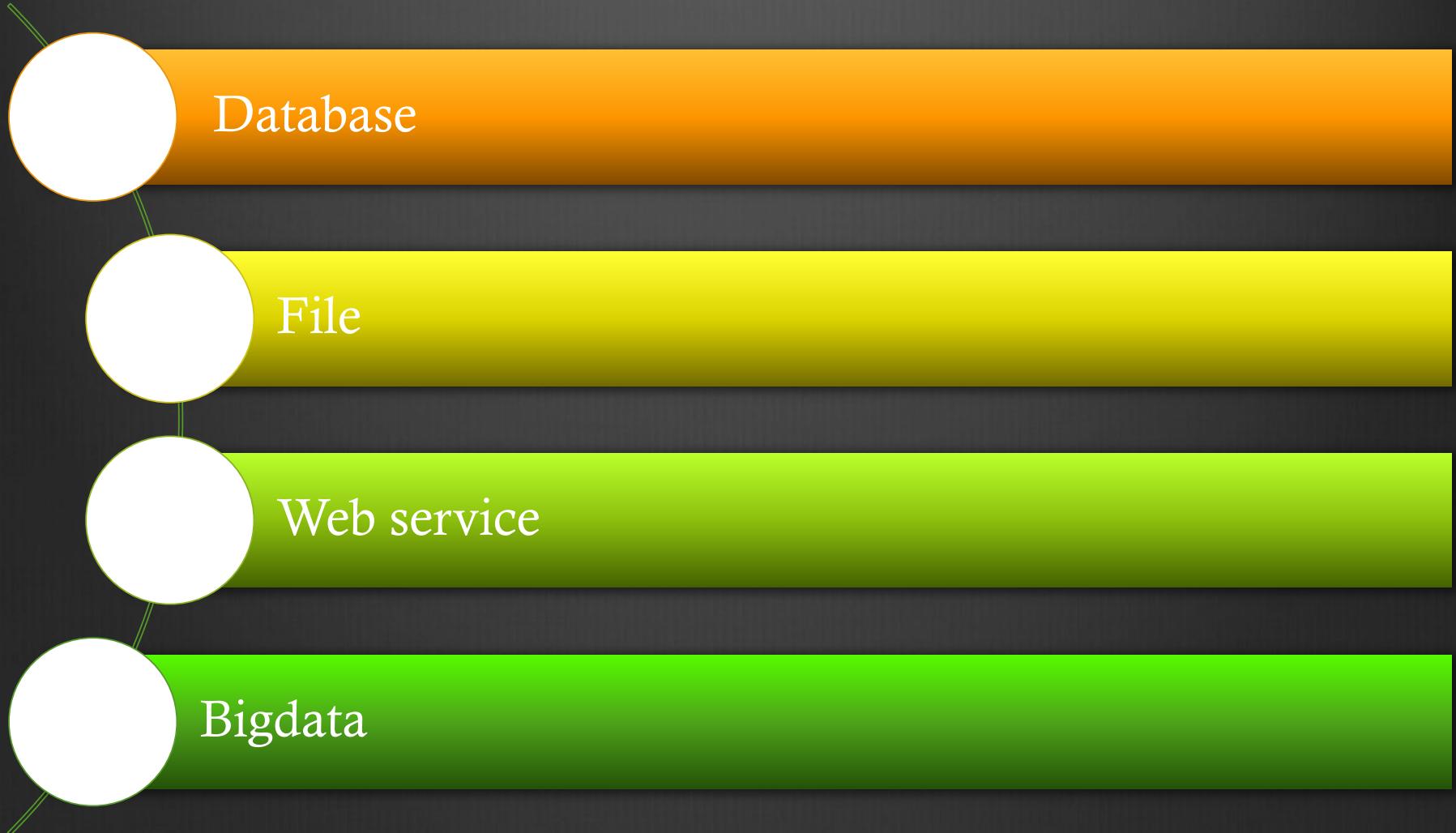
- Open-source
- Free
- Big community
- Easy to use
- Easy to debug



What is Kettle?

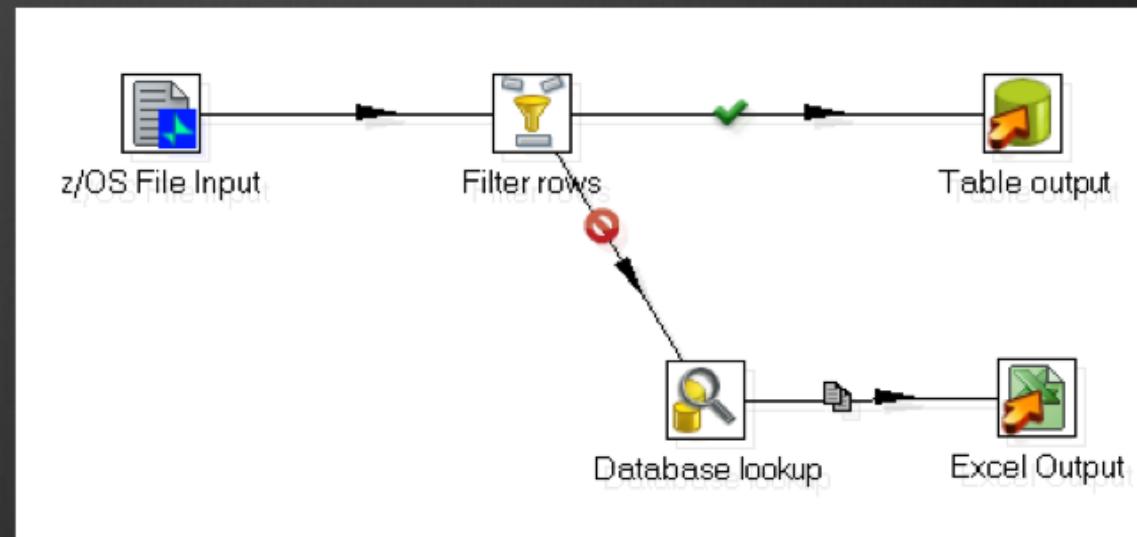
- Kettle – Pentaho Data Integration tool with powerful Extraction, Transformation, and Loading (ETL) capabilities
- Intuitive, graphical, drag and drop design environment
- Scalable and standards-based architecture

When to use



How it works

- Extraction (35+ databases, LDAP, files, web services, Bigdata, xBase, Geo-data, generated data ..)
- Transportation (union, append, filter, ..)
- Transformation (split, join, lookup, calculator, scripting, ...)
- Loading (see extraction)
- Streaming!!!

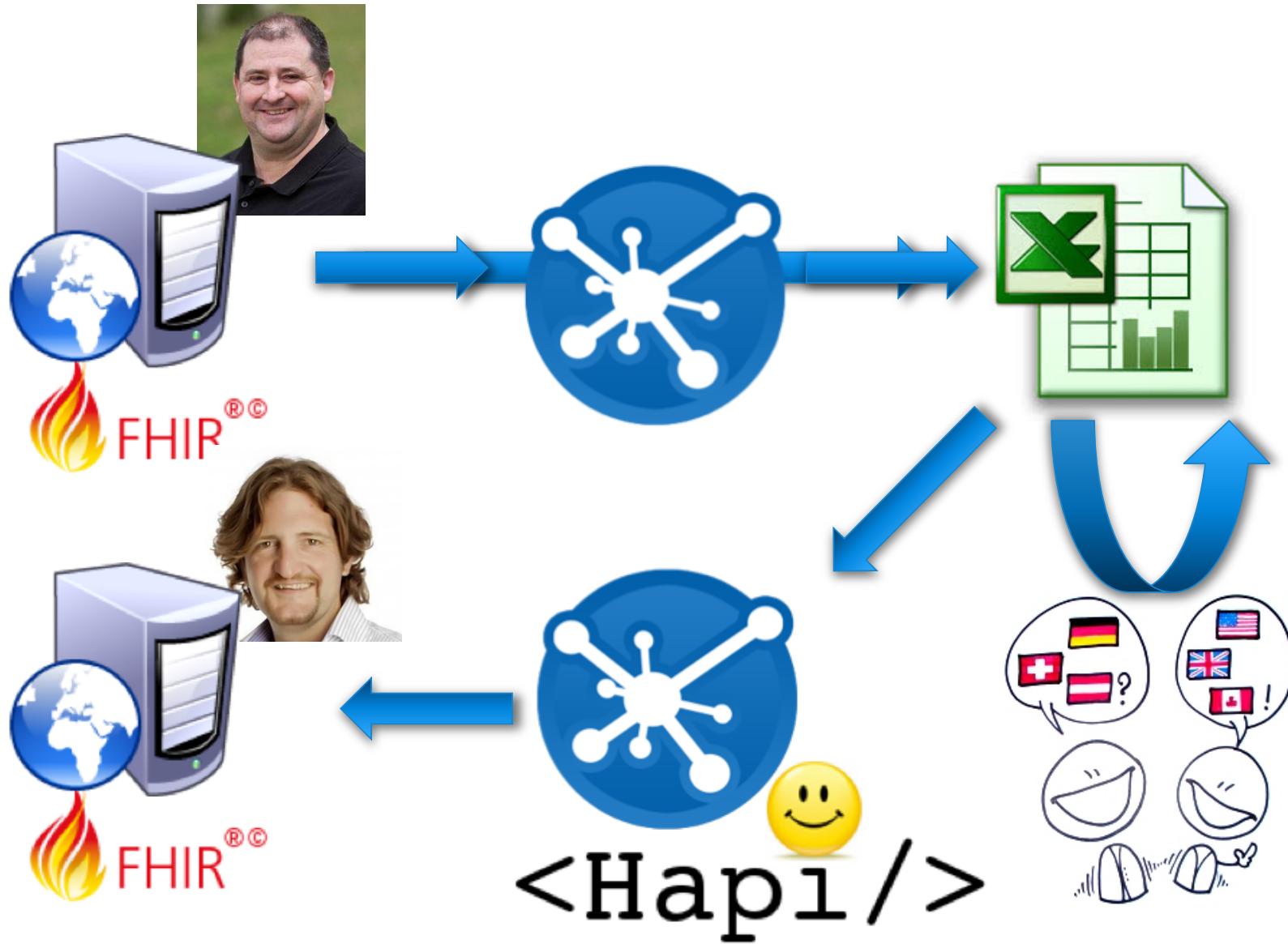


Examples

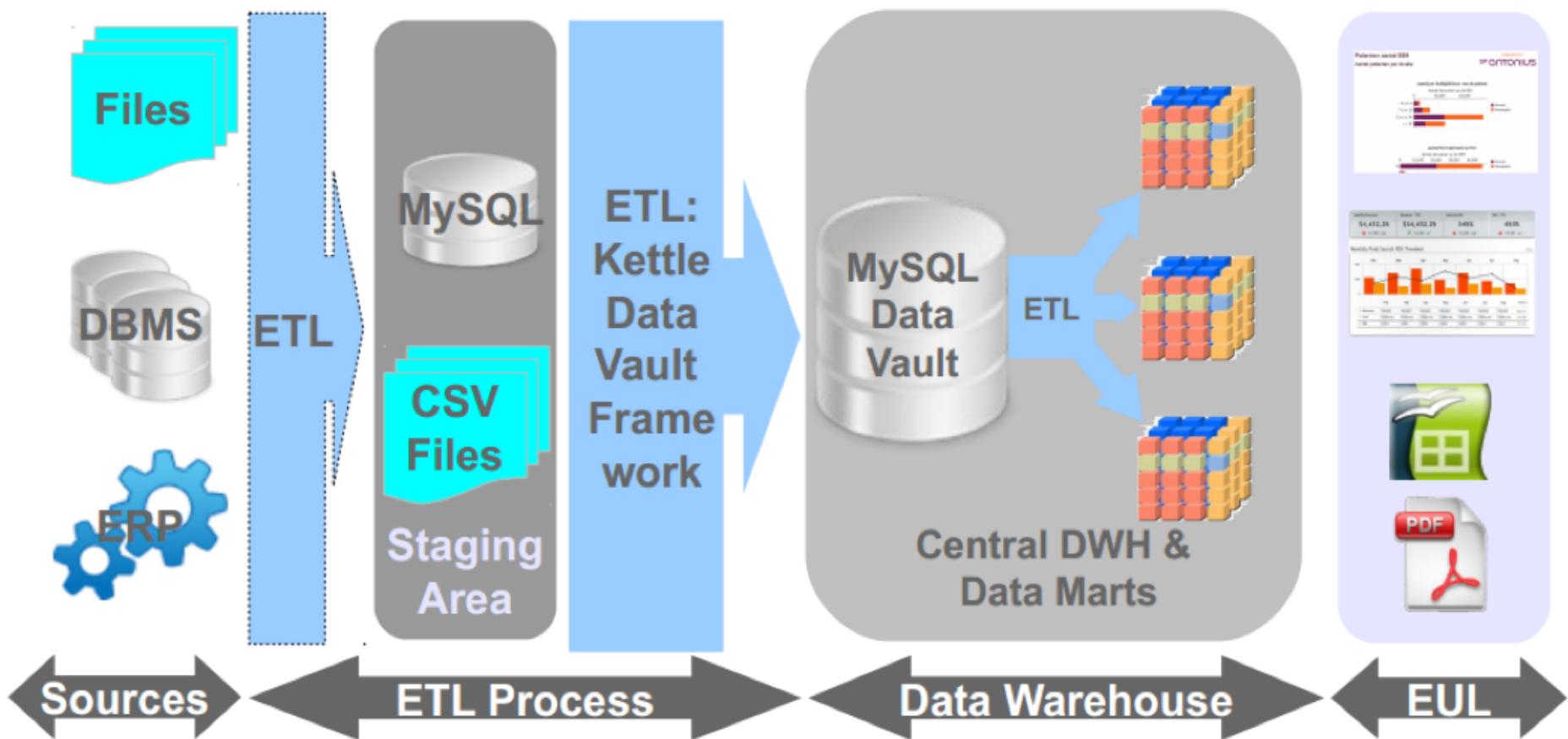
- Database -> Database
- Database -> Excel
- Javascript transformation (Period dimension)
- REST service <-> Excel
- Translation example (Property -> Excel)
- Overview of Nortal BI



ValueSet translations



Complex solution with Kettle



How to start

- Download and extract
zip<http://community.pentaho.com/projects/data-integration/>
- Run graphical designer - spoon.sh (.bat)
- Use kitchen.sh (.bat) to execute the jobs in a batch mode in shell (usually using scheduler)

Recommendation

Matt Casters

Chief of Data Integration
in Pentaho

Lead Development

Kettle Project Founder



Matt Casters
Roland Bouman
Jos van Dongen

Pentaho Kettle Solutions

Building Open Source ETL Solutions
with Pentaho Data Integration

A dynamic, blurred graphic at the bottom of the book cover, composed of streaks of orange, red, and yellow light, suggesting speed and data flow.

My translation tool

<https://github.com/nortal/fhir-etl>

Q&A

Thank you for listening!