From OSIsoft PI to Big Data Analytics: A data driven solution to reduce the environmental impact of upstream operations

Lorenzo Lancia & Gianmarco Rossi

Authors:

M. Montini, L. Cadei, G. Rossi, D. Loffreno L. Lancia, A. Corneo, D. Milana, M. Carrettoni, F. Landi, M. Galante, C. Bottani, V. Fostini

Agenda

- Project Scope
- Energy Efficiency analytics: eDea
- PI Connection: Data Science Lab
- PI Connection: Live Architecture
- Field Application
- Conclusions & Further Developments

Project Scope

Forecast & detect anomaly in energy consumption **exploiting real-time data** to:

- reduce energy consumption and CO_2 emissions from stationary combustion
- enhance hydrocarbon production
- improve asset integrity, process parameter optimization, HSE sustainability

Developing an integrated and real time data solution to:

Monitor and forecast global energy efficiency KPI

Promptly detect anomaly in energy consumption or efficiency

Help techincians to drill down into the root causes to find corrective actions

Reducing energy consumption is a key corporate objective:

low-carbon by reducing CO_2 emissions as a step toward carbon neutrality

decrease the environmental impact

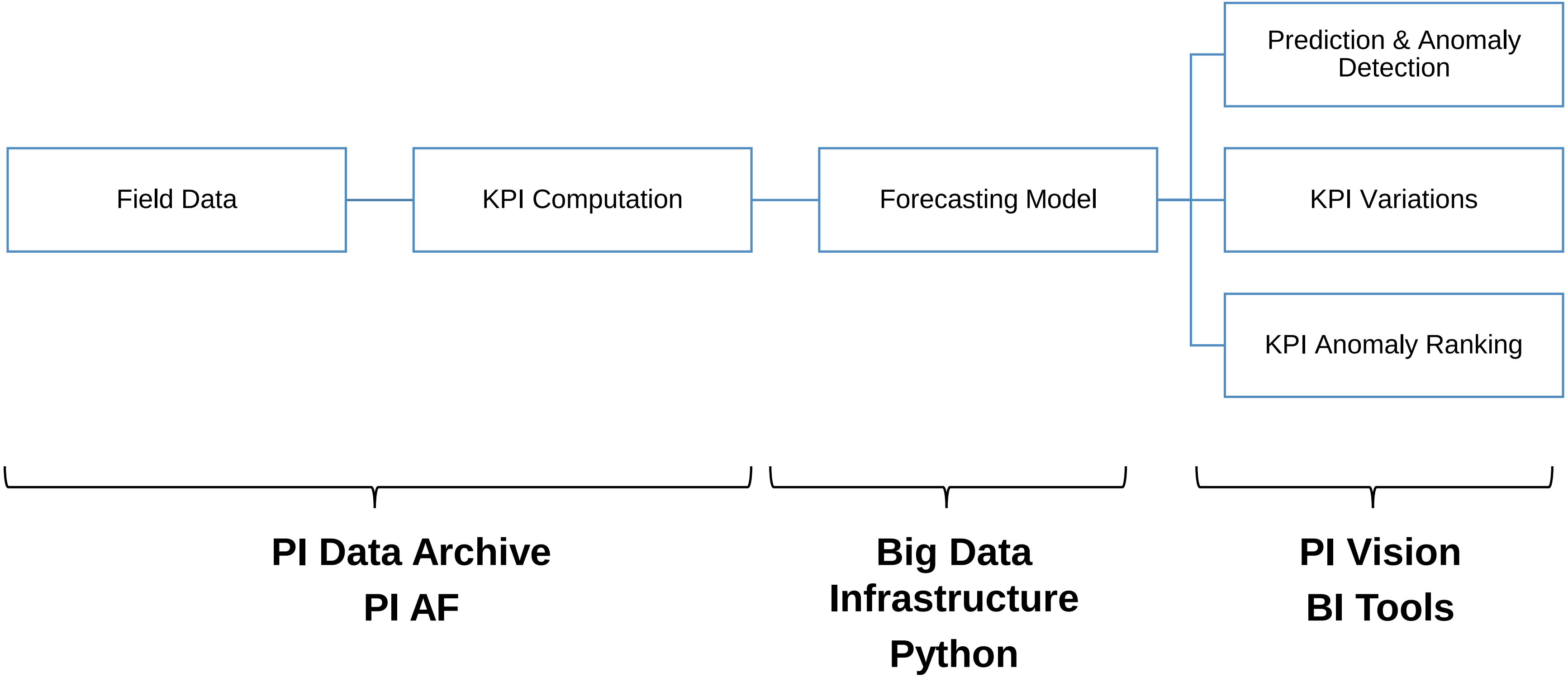
maximize hydrocarbon production by increasing efficiency

e-dea^{tm}

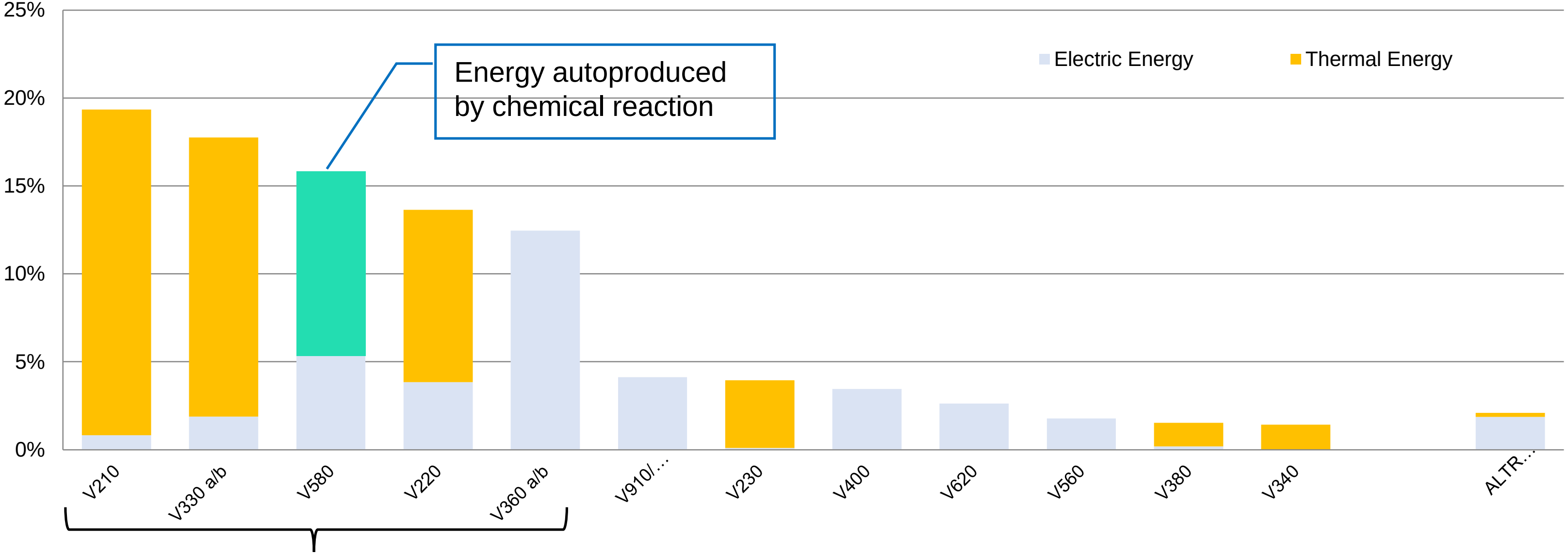
e-dea^{tm} is the analytics dashboard tool that leverages machine learning models to:

- Monitor & forecast the energy efficiency of an upstream plant
- Help technicians detect anomalies and suggest corrective actions

The e-dea^{tm} tool standard workflow



Energy consumption in an upstream plant is localized in key equipment:

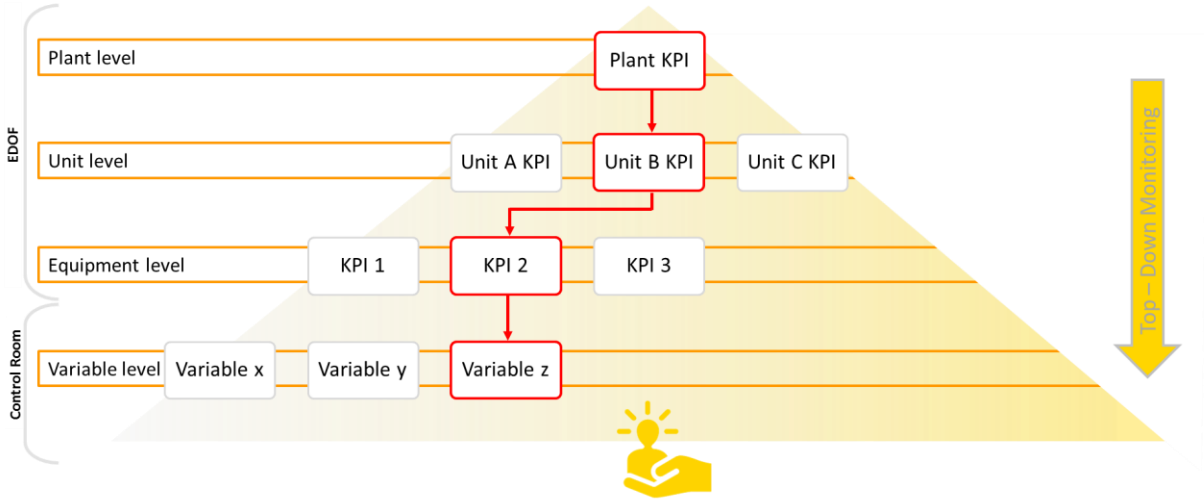


80% of Total consumption

Stationary Combustion CO_2 Emission (EI) is the main KPI to monitor:

$$EI = \frac{Fuel\ Gas\ \times Emission\ Factor}{Gross\ Hydrocarbon\ Production}\left[\frac{tCO_2}{kboe}\right]$$

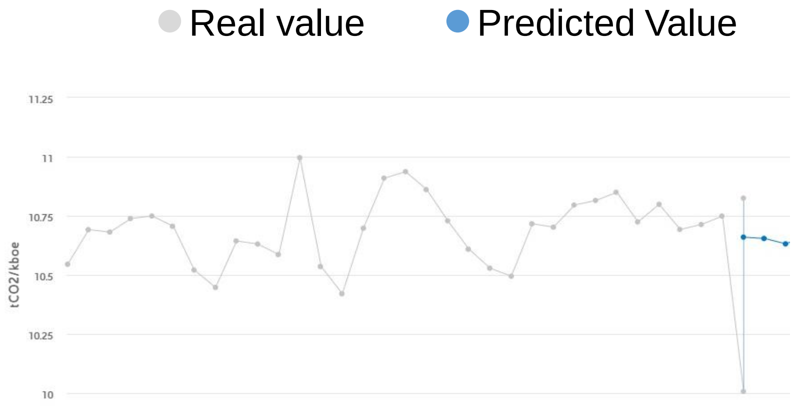
For each energy intensive equipment there are specific energy related KPIs



Forecasting Model

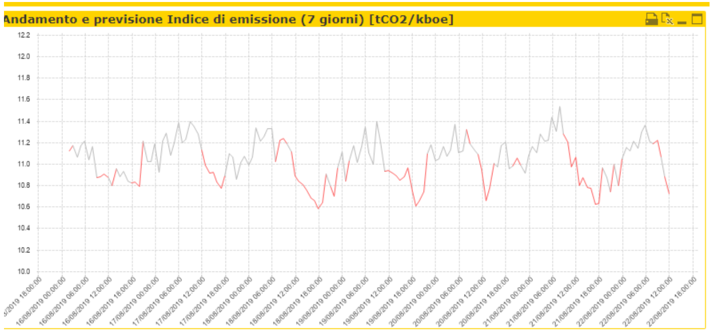
- Gradient Boosting Regression algorithm, predicts the value Stationary Combustion CO_2 Emission Index KPI for the next 3 hours.
- Predictors features are KPI and operational parameters for all the energy intensive equipment, seasonal features and exogenous like temperature or humidity.
- On the train/ test test the model achieved a R^2 about 0.80 and a MAPE about 5%.

STATIONARY COMBUSTION CO_2 EMISSION (EI)



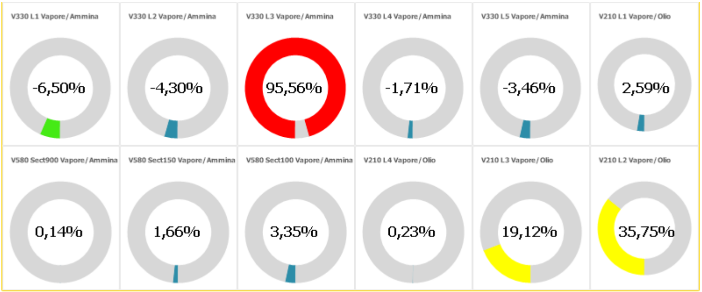
Prediction & Anomaly Detection

Using the predicted values, site operators will get a plant status report. By confronting real values with predictions we can detect anomalies in plant consumption.



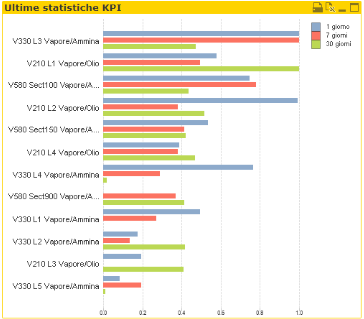
KPI Variations

In the event of an anomalous situation the dashboard can be used to check various equipment. Gauge graph show the variation of energy related KPIs with respect to different time frame.

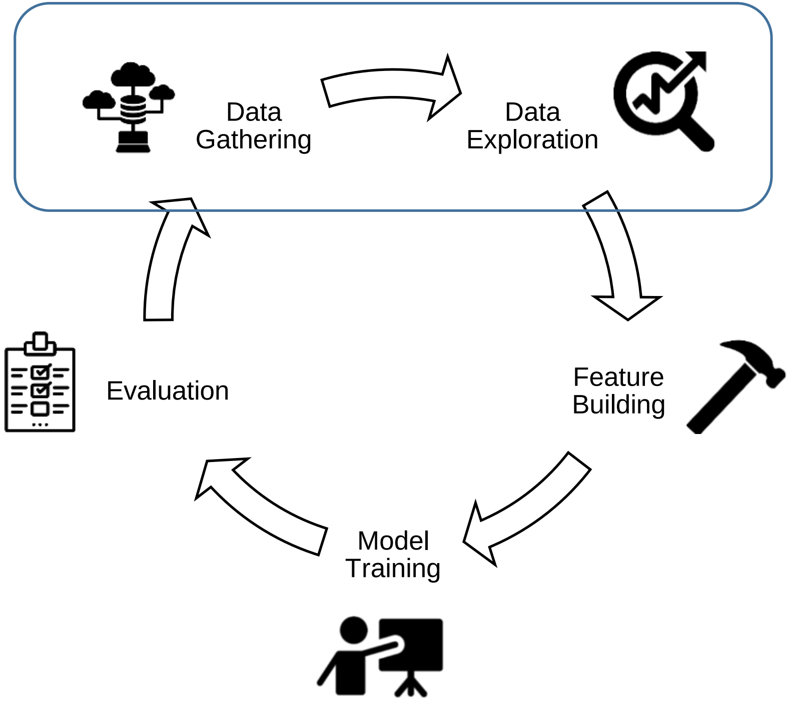


KPI Anomaly Ranking

To avoid temporary or non relevant fluctuations to trigger unwanted response. Another graph is shown in the dashboard displaying fluctuations normalized by sensor standard deviation



Developing a machine learning model means iterating through a series of steps:



And having to manage and organize data from different sources, deal with missing or not valid data.

Data Science development in Eni uses open source tools from the python environment.

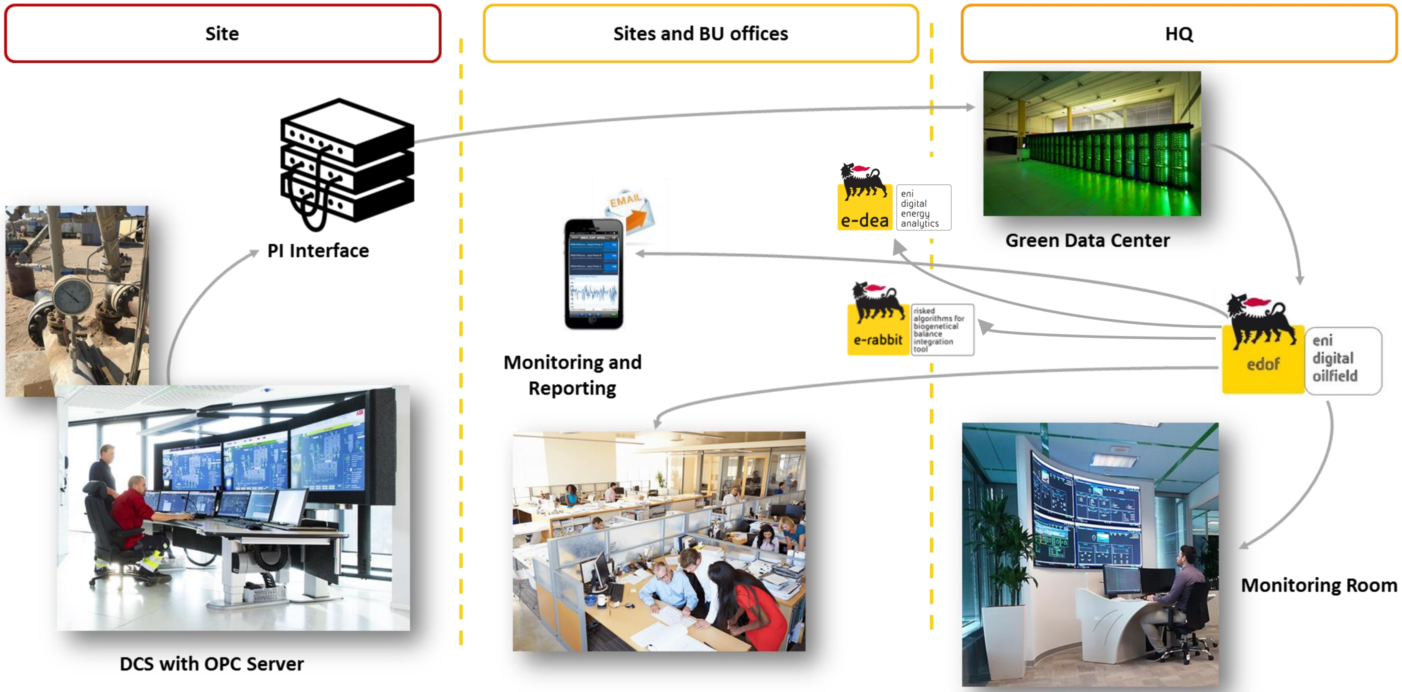
Source of Data – eDOF

- *Mature technology adopted by O&G Industry various names: Smart Fields®, Field of the Future®, i-Field®, Intelligent Field or Integrated Operations;*
- *Based on high frequency data acquired automatically in real time, integrated with lower frequency data (daily, monthly…), results of modelling and simulation and manually collected data to support better decision making processes;*

Eni standard configuration is named eDOF:

- based on configuration of state-of-the art off-the-shelf components
- design, implementation and deployment activities are performed by **Eni people**, both from IT and business disciplines, incorporating Eni Intellectual Property.
- eDOF is actually acquiring/calculating $350*10^6$ values every day, with an average frequency of 20 sec**.**

PI Connection – Standard eDOF infrastructure



Data from the asset are historicized at HQ Green Data Center

Development of the system from HQ in tight collaboration with BU engineers

Full support from Eni HQ

PI speeding up the data science workflow

Fast and flexible access sensors time series

Consistent Data aggregation, Interpolation and KPIs Computation

Zero *missing at random* data

3 steps to train a model from PI System data

Explore source data to identify relevant time series

- Autonomous data gathering

Ingest only relevant time series in Big Data environment

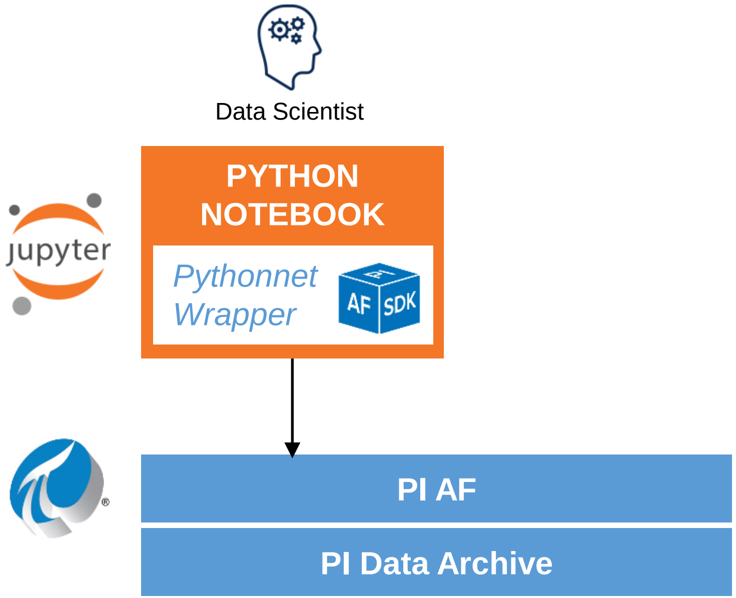
- Automatic ingestion
- Consistency in data provided to modelling phase
- Continuous data update

Start modelling with data from big data storage

- Coherence in the environments and used in development
- and production machine pipeline.
- Reduce query workload on critical systems.

Explore source data

Data discovery performed via direct connection to PI from Data Science development environment



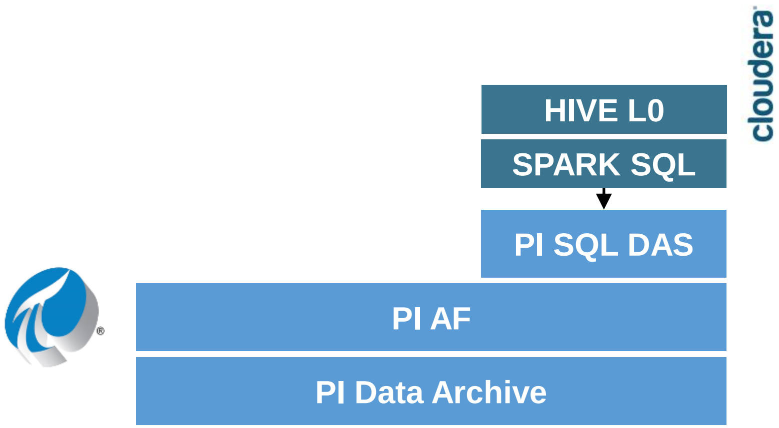
ADVANTAGES

- Zero configuration access to PI Data Archive
- Complete access to time series and PI functionalities
- Security granted by PI profiles and NT authentication
- Language and environment optimized for data science
- Real-time update

```python
def get_tag_history(self, tag_name, start_date, end_date, delta_t, convert_to_string=True):

    tag_obj = PIPoint.FindPIPoint(self.piServer, tag_name)
    str_startdate = str(start_date)
    str_enddate = str(end_date)

    timerange = AFTimeRange(str_startdate, str_enddate)
    span = AFTimeSpan.Parse(delta_t)

    interpolated = tag_obj.InterpolatedValues(timerange, span, "", False)
```

Ingest relevant time series

Big data platform ingests relevant time series identified during data discovery activity and make them available to production AI models for prediction and forecast.

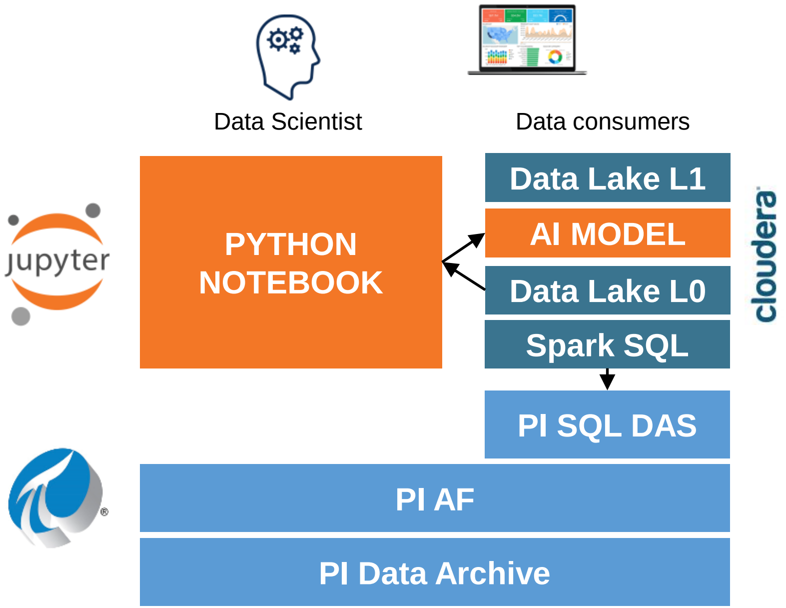


ADVANTAGES

- No need to perform utopic PI complete data ingestion
- Data updated every 5 minutes for relevant time series
- Environment optimized for algorithm execution
- Reduced and efficient workload on PI server
- Data structure optimized fro machine learning pipelines

Start modelling with data

Modelling phase works with ingested official data; data scientists development is directly integrated with Big Data environment enabling data access and AI model deployment

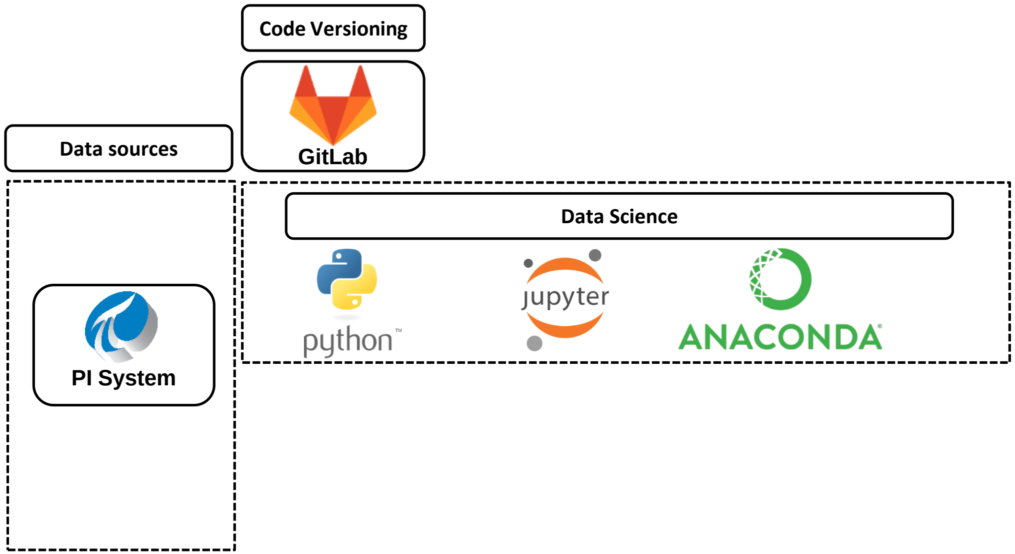


ADVANTAGES

- Common official data sets to all AI models
- Same data used for development and pipelines
- Access granted to pipeline logs and outputs
- Data scientists work in the same environment used for data discovery
- Direct model deployment to productive pipelines
- Output data exposed to dashboards and applications

Technological Stack

Exploration phase

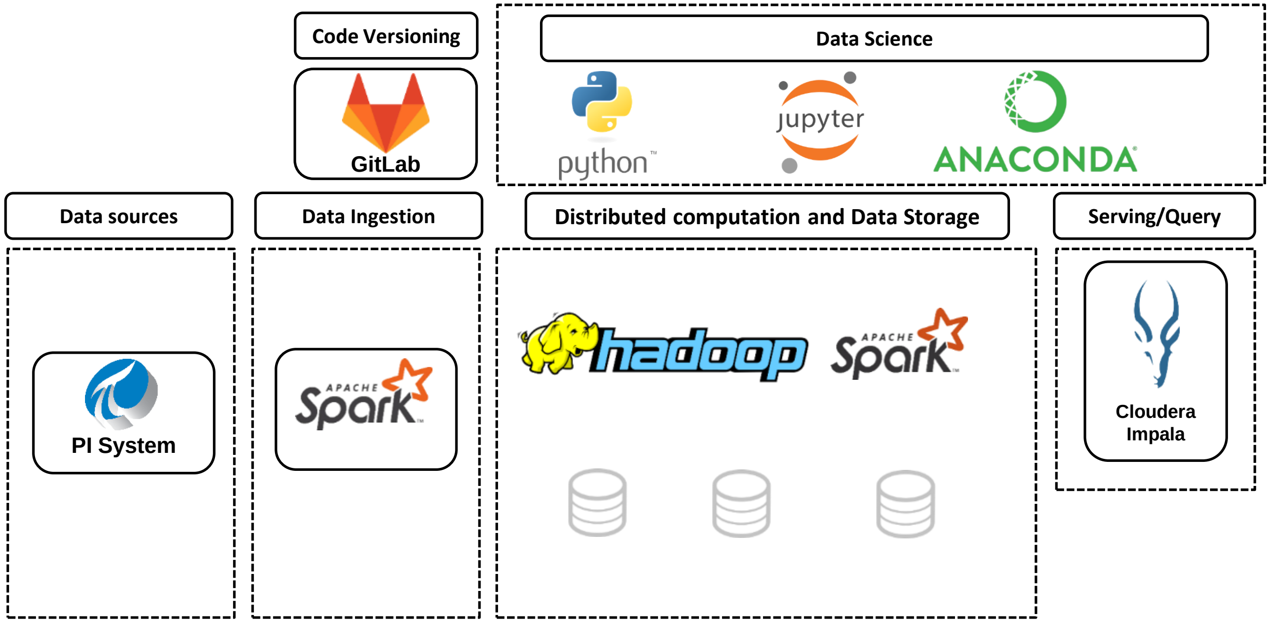


Modelling

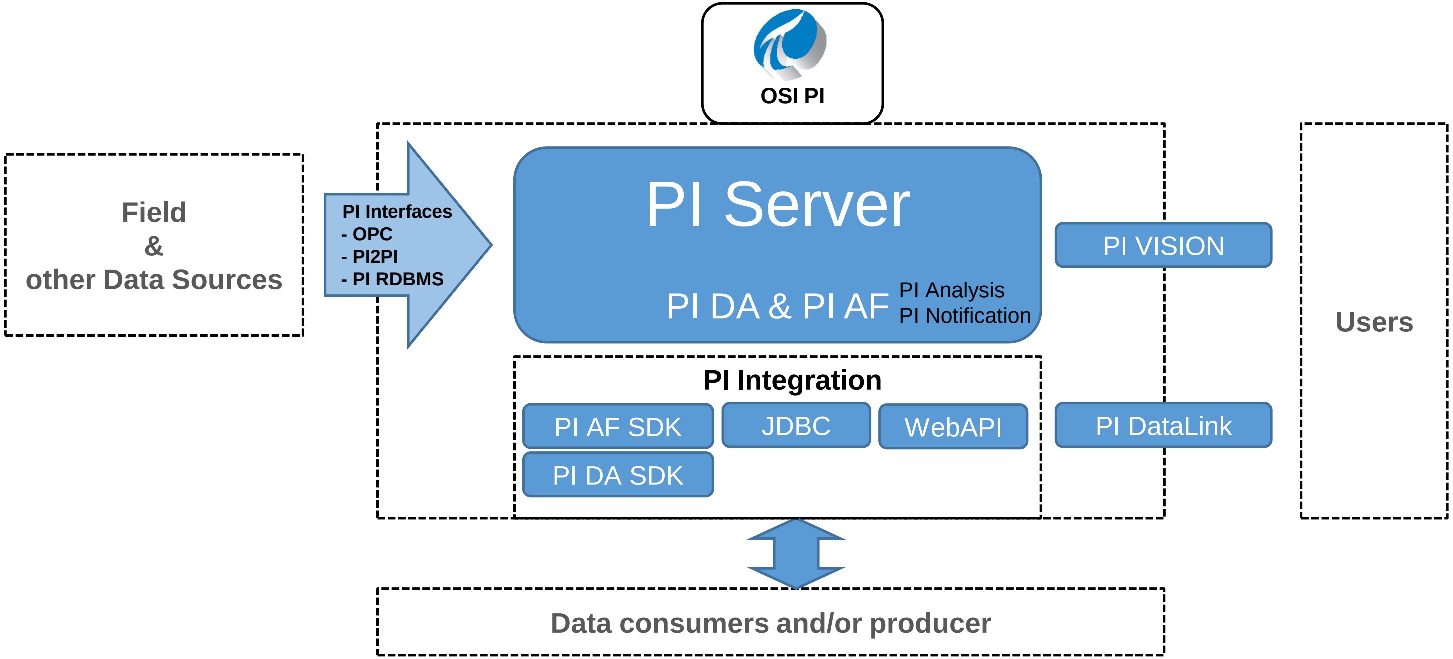
- Training a ML Regressor with tag time series batch collected from Big Data
- Only essential feature transformation demanded to model. Most operation assigned to either PI Server or Big Data Spark Job.
- Training done in Jupyter Environment to evaluate performances.
- After training a serialized object (pickle) is exported for deploying the model for live scoring.

Technological Stack

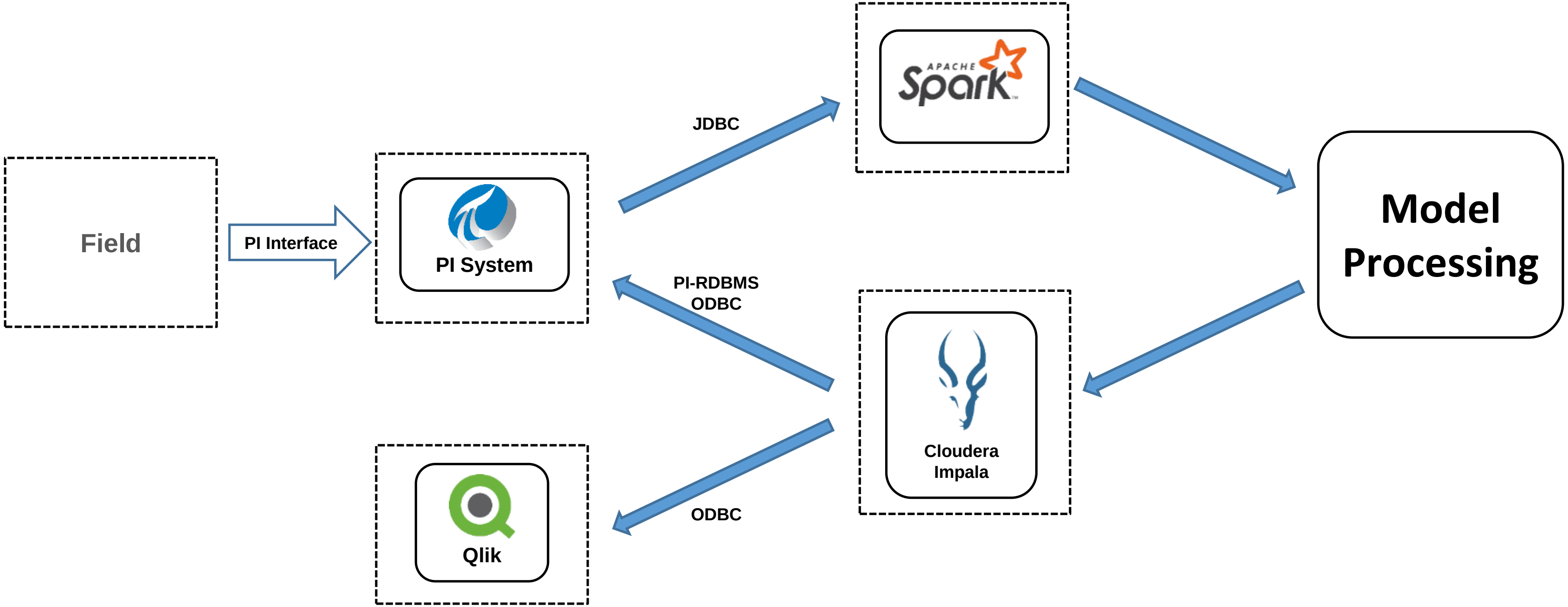
Modelling phase



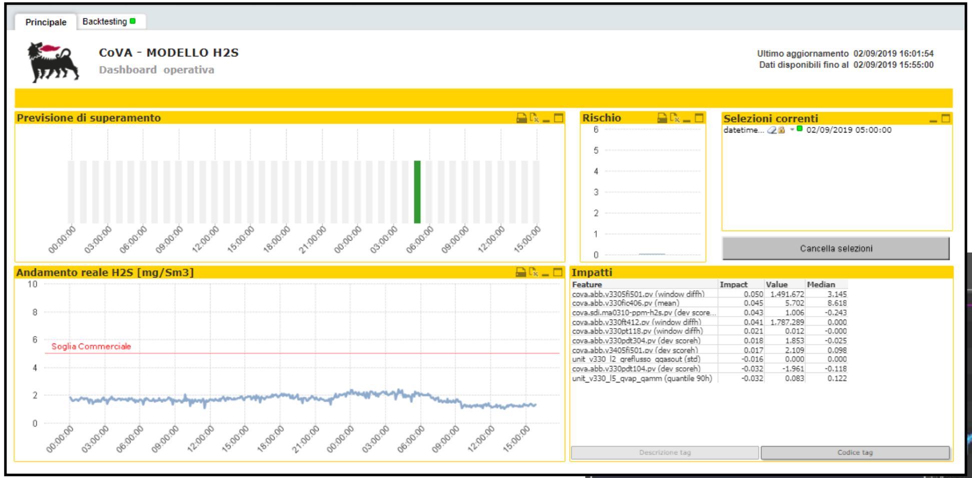
PI System Architecture



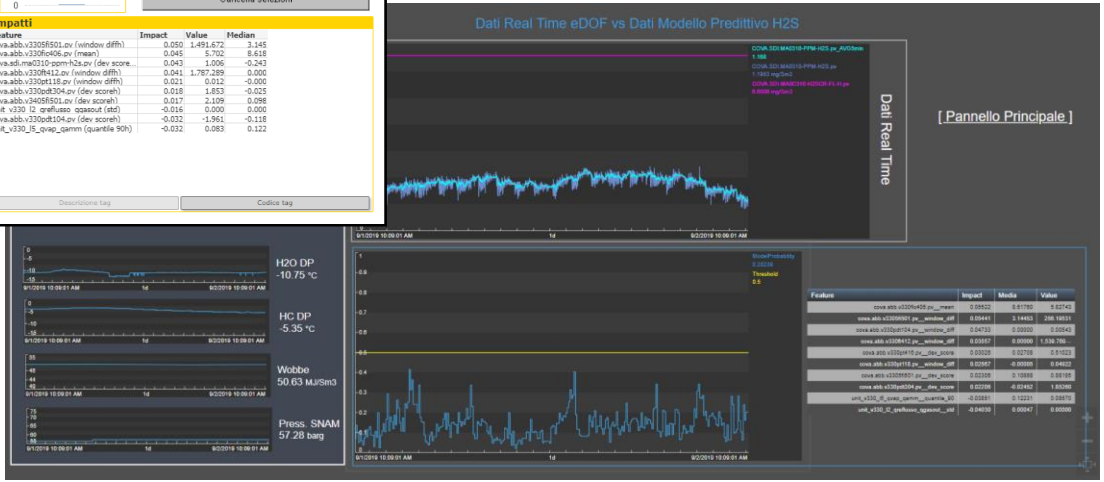
Data Flow and Connection Architecture



Data availability: BI Tools and PI Vision

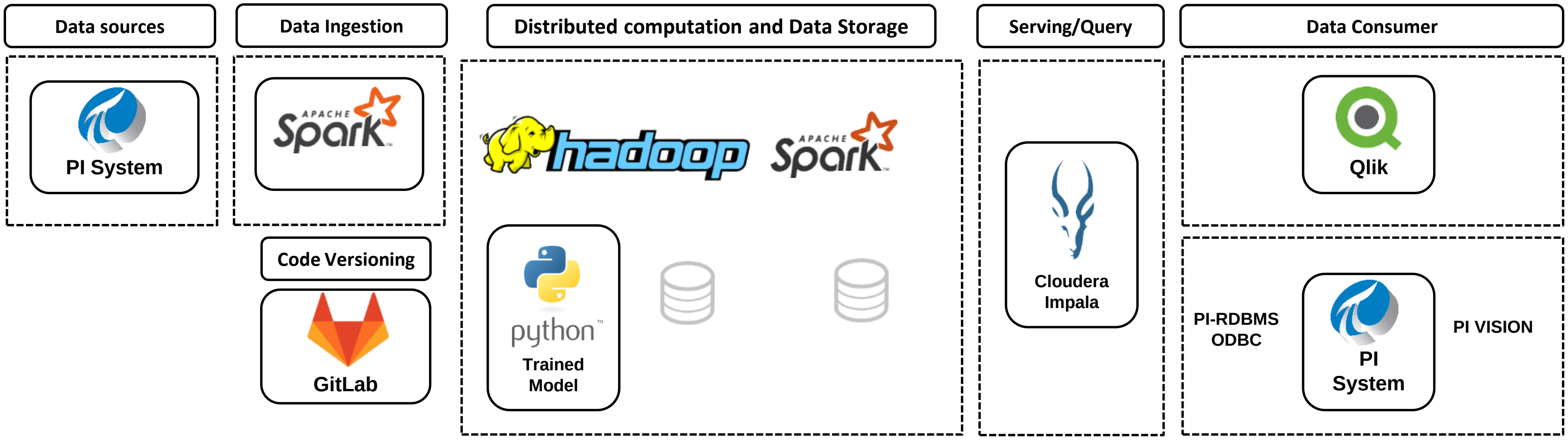


Feature	Impact	Value	Median
cova.abb.v3305fi501.pv (window diffh)	0.050	1.491.672	3.145
cova.abb.v330fic406.pv (mean)	0.045	5.702	8.618
cova.sdi.ma0310-ppm-h2s.pv (dev score...	0.043	1.006	-0.243
cova.abb.v330ft412.pv (window diffh)	0.041	1.787.289	0.000
cova.abb.v330pt118.pv (window diffh)	0.021	0.012	-0.000
cova.abb.v330pdt304.pv (dev scoreh)	0.018	1.853	-0.025
cova.abb.v3405fi501.pv (dev scoreh)	0.017	2.109	0.098
unit_v330_l2_qreflusso_qgasout (std)	-0.016	0.000	0.000
cova.abb.v330pdt104.pv (dev scoreh)	-0.032	-1.961	-0.118
unit_v330_l5_qvap_qamm (quantile 90h)	-0.032	0.083	0.122



Technological Stack

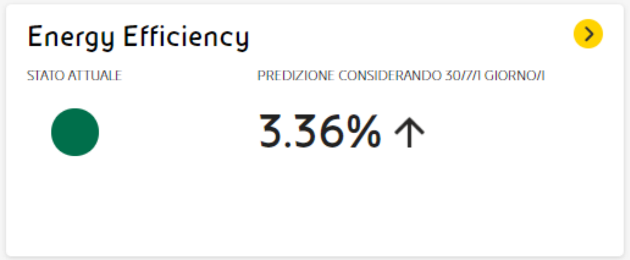
Live Scoring Phase

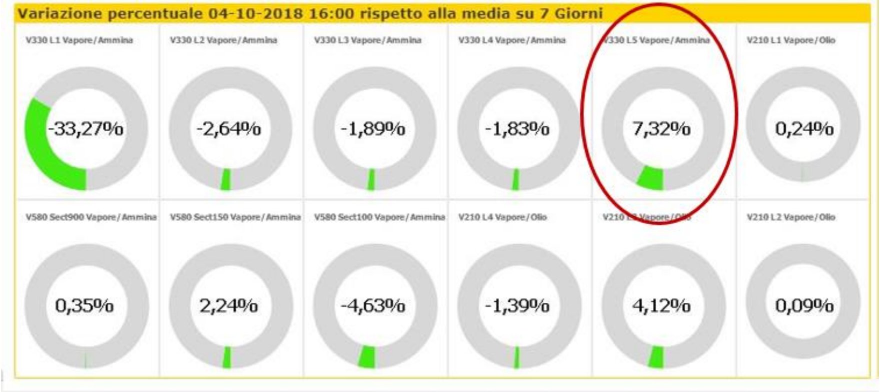


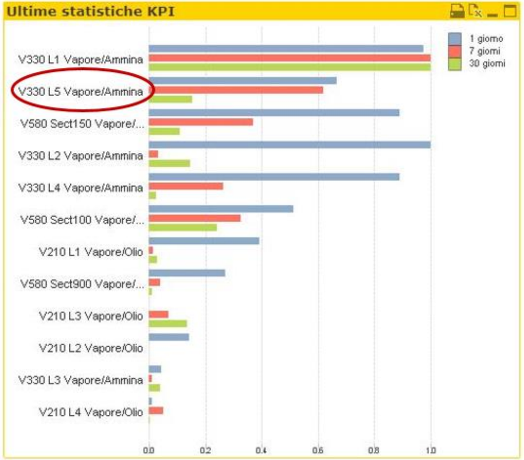
All product names, trademarks and registered trademarks are property of their respective owners.

Field application: real case

- e-dea^tm detects an anomaly in energy consumption, foreseeing the increasing of the KPI
- and drill down into the root causes, indicating the equipment with bad performance (positive variation %)







Field application: Real case

- The production engineer easily checks the parameters and trends on PI-Vision, knowing in advantage on which unit to look.
- Action to reduce energy consumption are implemented and monitored.
- The tool is also very useful to restore the optimal condition after a variation in the operating conditions.

Data Driven Energy Efficiency

CHALLENGES

- How to get access to PI Data from Data Science standard tools.
- Deploy a Model into Big Data environment
- Make the output of data science model available in tools familiar to operations technicians like PI VISION

SOLUTION

- Use Pythonnet to wrap PI AFSDK
- Use JDBC connection to PI to ingest data into Big Data
- Write back output into PI from Impala via ODBC

BENEFITS

- Flexible access during data exploration phase
- Structured access to PI AF & PI Data Archive
- Discovery of possible efficiency actions by leveraging on machine learning tools

e-dea

eni digital energy analytics

> “From the beginning of real field application we implemented more than 15 energy efficiency actions leading to a significant reduction in co2 emission of an upstream giant oil field”

Speakers

- Lorenzo Lancia
- lorenzo.lancia@eni.com
- Data Scientist
- Eni

- Gianmarco Rossi
- Gianmarco.rossi@eni.com
- Production Engineer
- Eni

Questions?

Please wait for the **microphone**

State your **name & company**

Please remember to…

Complete Survey!

Navigate to this session in mobile agenda for survey

謝謝 KEA LEBOHA
DZIĘKUJĘ CI
TAPADH LEIBH
고맙습니다
GRÀCIES
БАЯРЛАЛАА
MISAOTRA ANAO
SALAMAT
NGIYABONGA
OBRIGADO شكرا
TEŞEKKÜR EDERIM
KÖSZÖNÖM
DANKIE
TERIMA KASIH
DANKON TANK TAPADH LEAT
СПАСИБО
MULȚUMESC
OSIsoft PI World
FAAFETAI
РАКМЕТ СІЗГЕ
HVALA
ESKERRIK ASKO
GO RAIBH MAITH AGAT
MAHADSANID
HVALA ХВАЛА ВАМ
БЛАГОДАРЯ GRACIAS
THANK YOU
TEŞEKKÜR EDERIM
ТИ БЛАГОДАРАМ
DANK JE
ΕΥΧΑΡΙΣΤΩ GRATIAS TIBI
ДЗЯКУЙ
GRAZIE
TAK DANKE
AČIŪ SALAMAT MAHALO IĀ ʻOE TAKK SKAL DU HA
DI OU MÈSI
RAHMAT MERCI
GRAZZI ÞAKKA ÞÉR
FALEMINDERIT
ありがとうございました
ĎAKUJEM
HATUR NUHUN
РАХМАТ САҒА
SIPAS JI WERE TERIMA KASIH
MATUR NUWUN
CẢM ƠN BẠN
UA TSAUG RAU KOJ
WAZVIITA
ТИ БЛАГОДАРАМ
СИПОС