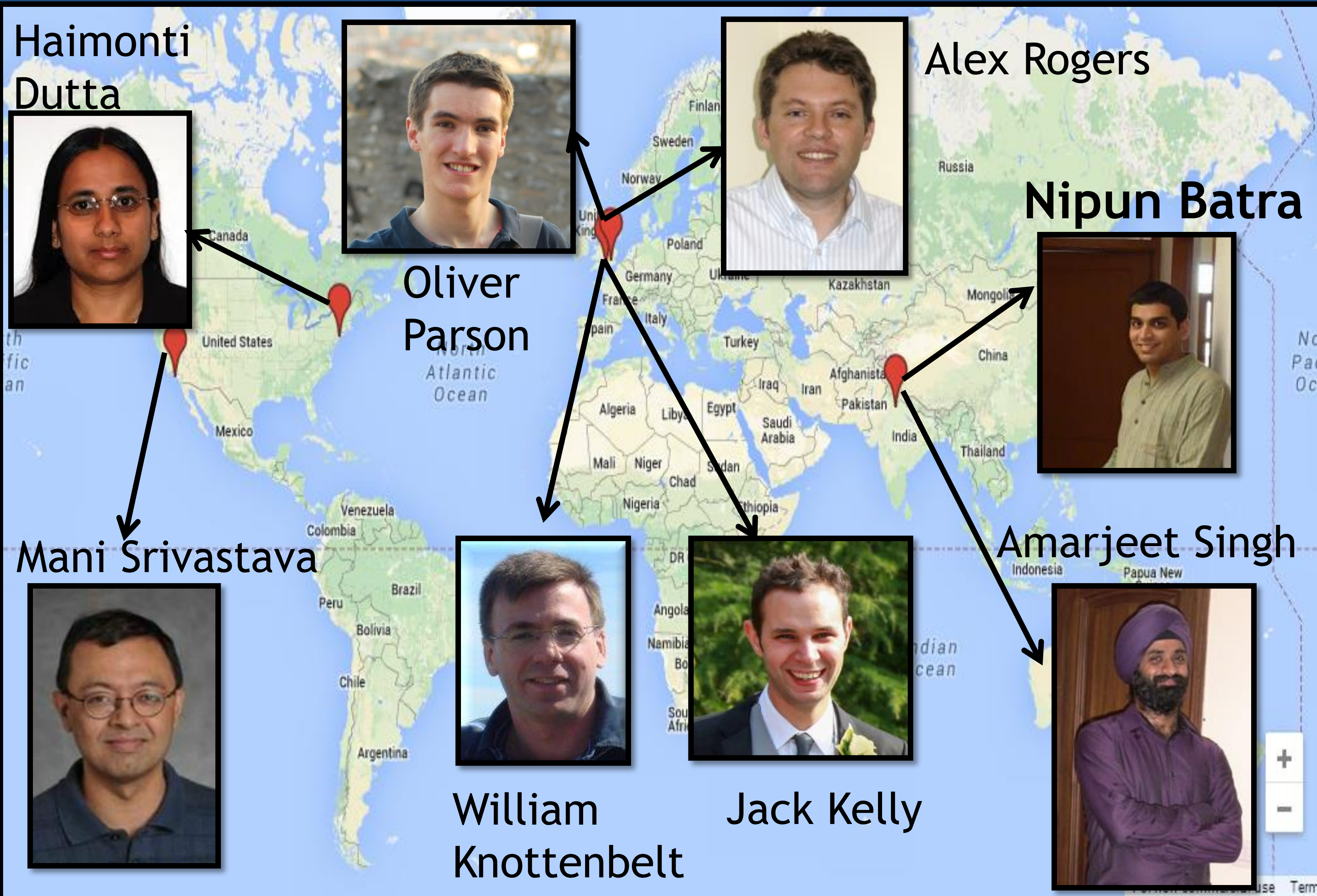
INDRAPRASTHA INSTITUTE of INFORMATION TECHNOLOGY DELHI
UCLA
COLUMBIA CCLS
Imperial College London
UNIVERSITY OF Southampton

NILMTK

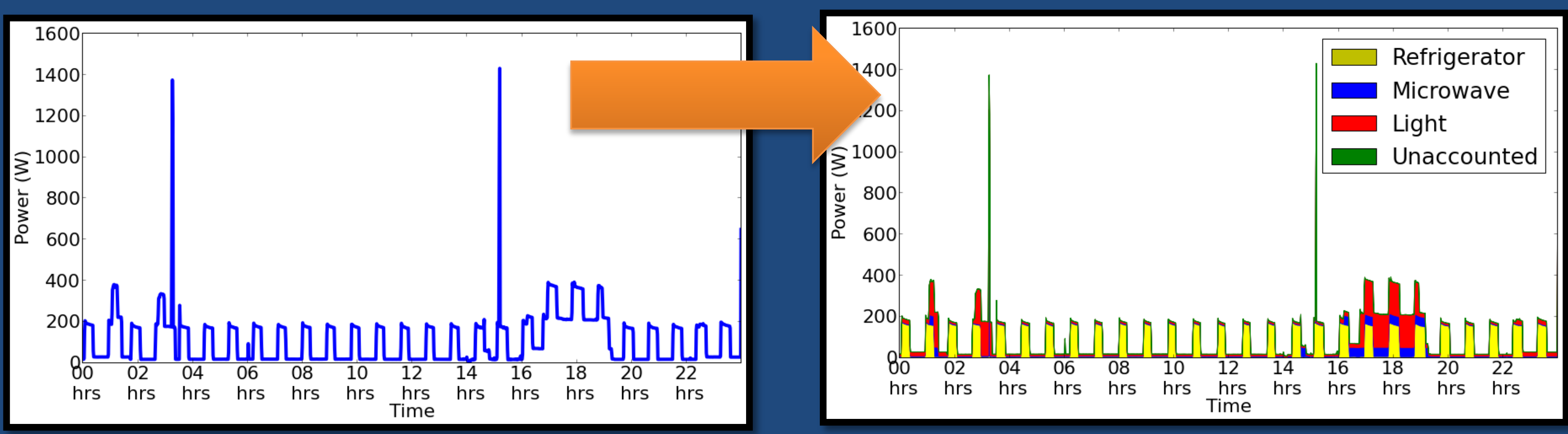
An Open Source Toolkit for Non-intrusive Load Monitoring

NILMTK team


Haimonti Dutta
Oliver Parson
Alex Rogers
Nipun Batra
Mani Srivastava
William Knottenbelt
Jack Kelly
Amarjeet Singh
Canada
United States
Mexico
North Atlantic Ocean
Sweden
Norway
Poland
Germany
Italy
Turkey
Russia
Kazakhstan
Mongolia
China
Iraq
Iran
Pakistan
India
Thailand
Algeria
Egypt
Saudi Arabia
Mali
Niger
Chad
Nigeria
Venezuela
Colombia
Brazil
Peru
Bolivia
Chile
Argentina
Angola
Namibia
Indonesia

Non-intrusive load monitoring (Energy disaggregation)

"Process of estimating the energy consumed by individual appliances given just a whole-house power meter reading"



Wait a minute! This sounds complicated Would it help?

Jane goes to the market

Jane spends 200 pounds on her purchases

Jane's husband John is worried with the expenses

He spends some time and looks at the purchase list

Do you think the itemized billing helped him?

NILM is the same, but for energy!

Quiz time!

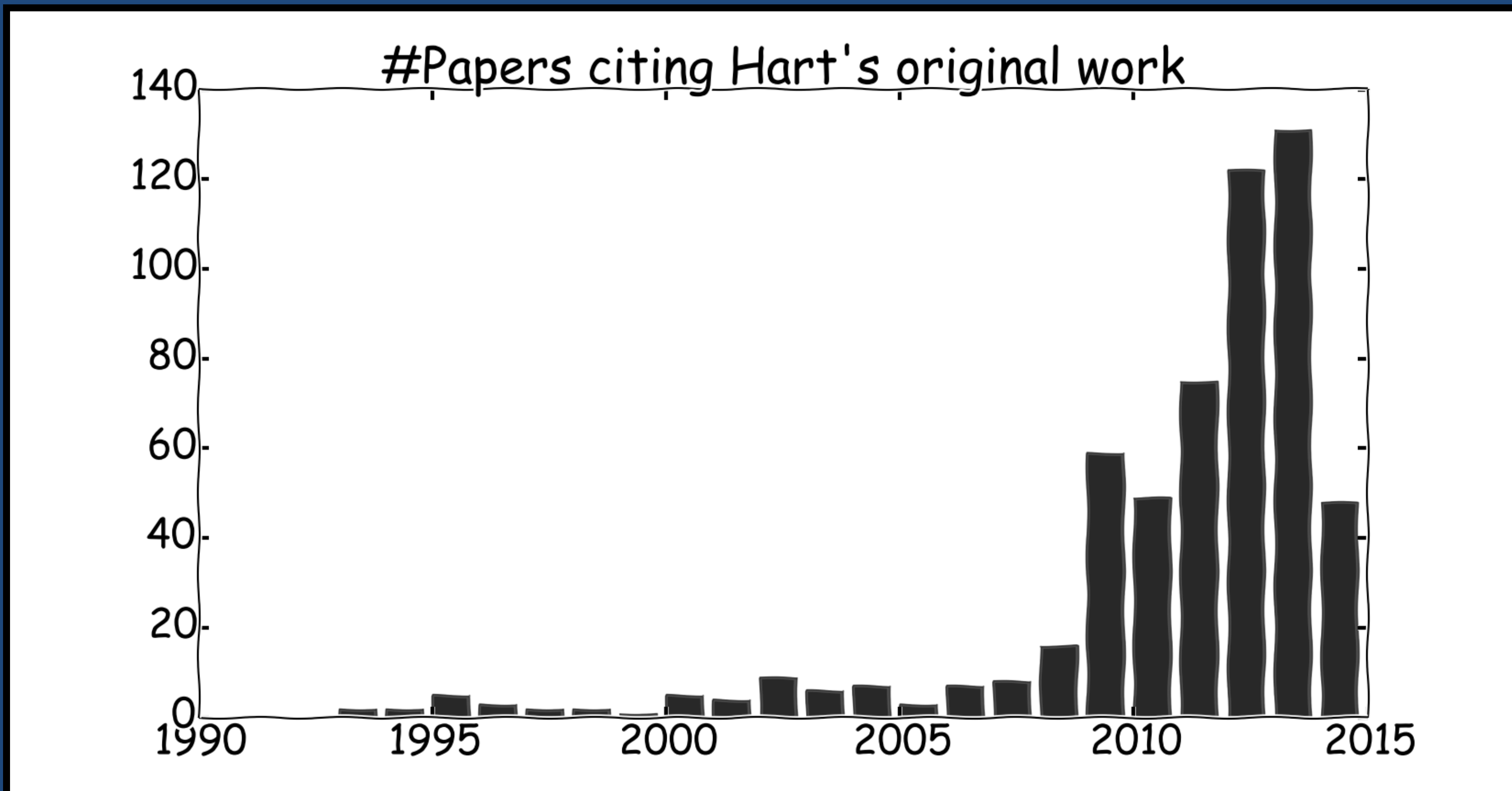
Identify this famous CS scientist

Quiz time!
Identify this famous CS scientist

That ain't any great scientist. That's me on my first birthday in 1990…

This is not too far from the time when NILM was first discussed

Giving credit where it is due



NILM interest explosion

1. National smart meter rollouts
2. Reduced hardware costs
3. International meetings
   - NILM workshop 2012, 2014; EPRI NILM 2013
4. Public datasets
5. Startups

“Data is the new oil”

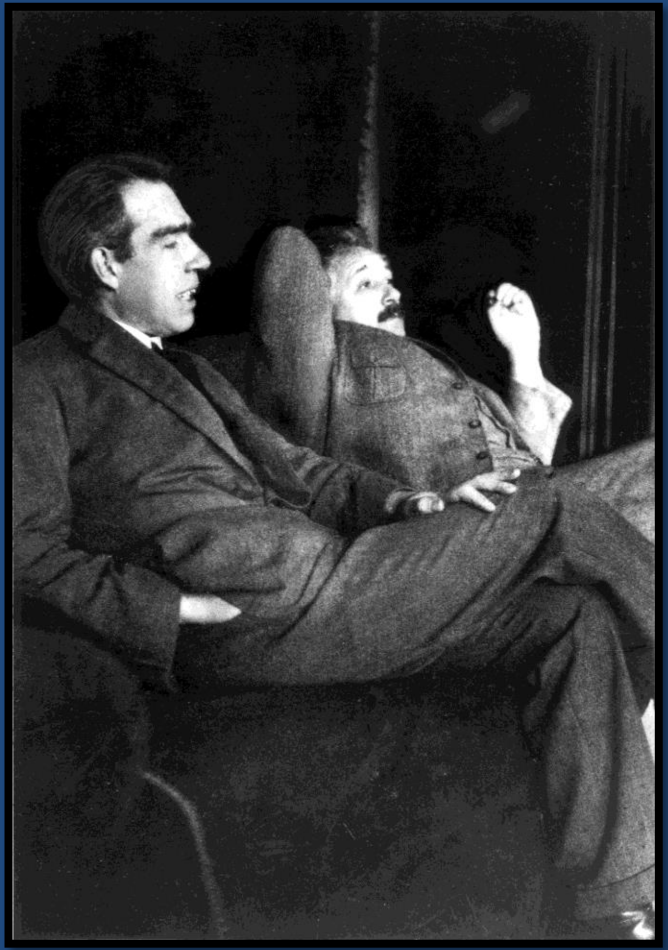
- 9 NILM datasets and counting (few not specific to NILM)
- Across 6 countries (India, UK, US, Canada, EU)
- Measure aggregate and appliance level data
- Across 3 colors ☺
  - REDD
  - BLUED
  - GREEND

The industry is interested!

So, is everything so rosy?
Not quite! Else we won’t be here

The scientific method

"The scientific method is a body of techniques for investigating phenomena, acquiring new knowledge, or **correcting and integrating previous knowledge**" as per wiki



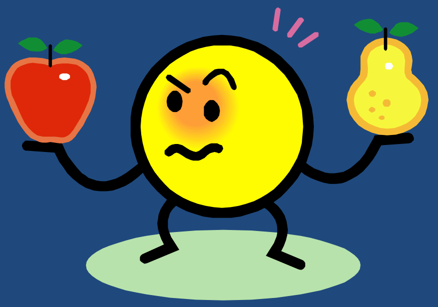
3 core obstacles preventing comparison of state-of-the-art

1. Hard to assess generality

- Subtle differences in aims of different data sets
- Previous contributions evaluated only on single dataset.
- Non-trivial to set up similar experimental conditions for direct comparison.

2. Lack of comparison against same benchmarks

- Newly proposed algorithms rarely compared against same benchmarks.
- Lack of “open source” reference algorithms → often lead to reimplementation.



3. "Inconsistent" disaggregation performance metrics

- Different performance metrics proposed in the past.
- Different formulae for same metric, eg. 4+ versions of "energy assigned"

$$\mathrm{Acc} = 1 - \frac{\sum_{t=1}^{T} \sum_{i=1}^{n} \left| \hat{y}_t^{(i)} - y_t^{(i)} \right|}{2 \sum_{t=1}^{T} \bar{y}_t}$$

$$\sqrt{\left(\sum_{t,i} \left\| y_t^{(i)} - \hat{y}_t^{(i)} \right\|_2^2 \right) / \left(\sum_{t,i} \left\| y_t^{(i)} \right\|_2^2 \right)}$$

$$\left| \sum_t x_t^{(n)} - \sum_t \mu_{z_t^{(n)}}^{(n)} \right| / \sum_t x_t^{(n)}$$

$$MNE(n) = \frac{\sum_{t=1}^{T} |\theta_t^n - y_t^n|}{\sum_{t=1}^{T} \theta_t^n}$$

What is NILMTK?

Open source NILM toolkit

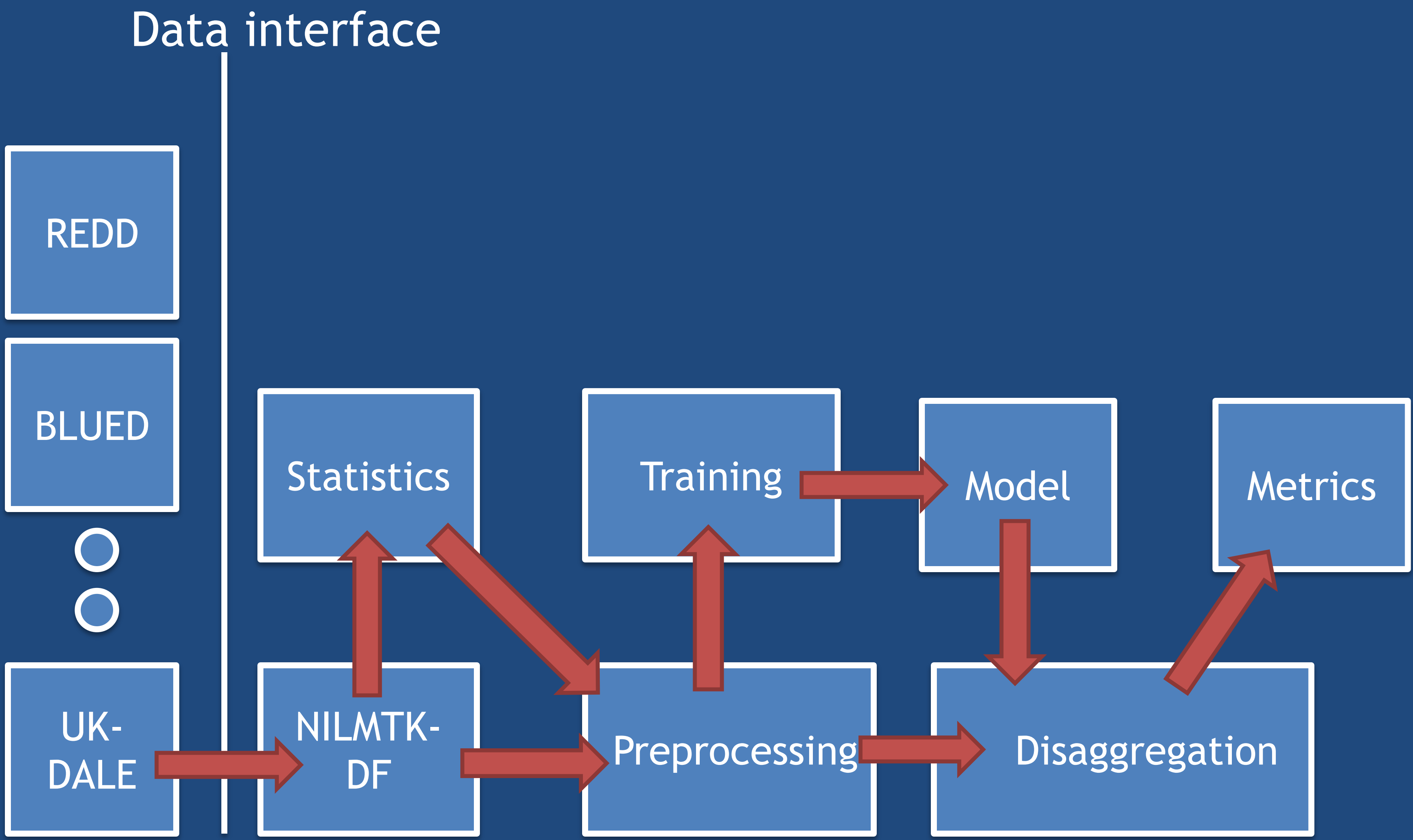
What does it do?

Enable easy comparative analysis of NILM algorithms across data sets.

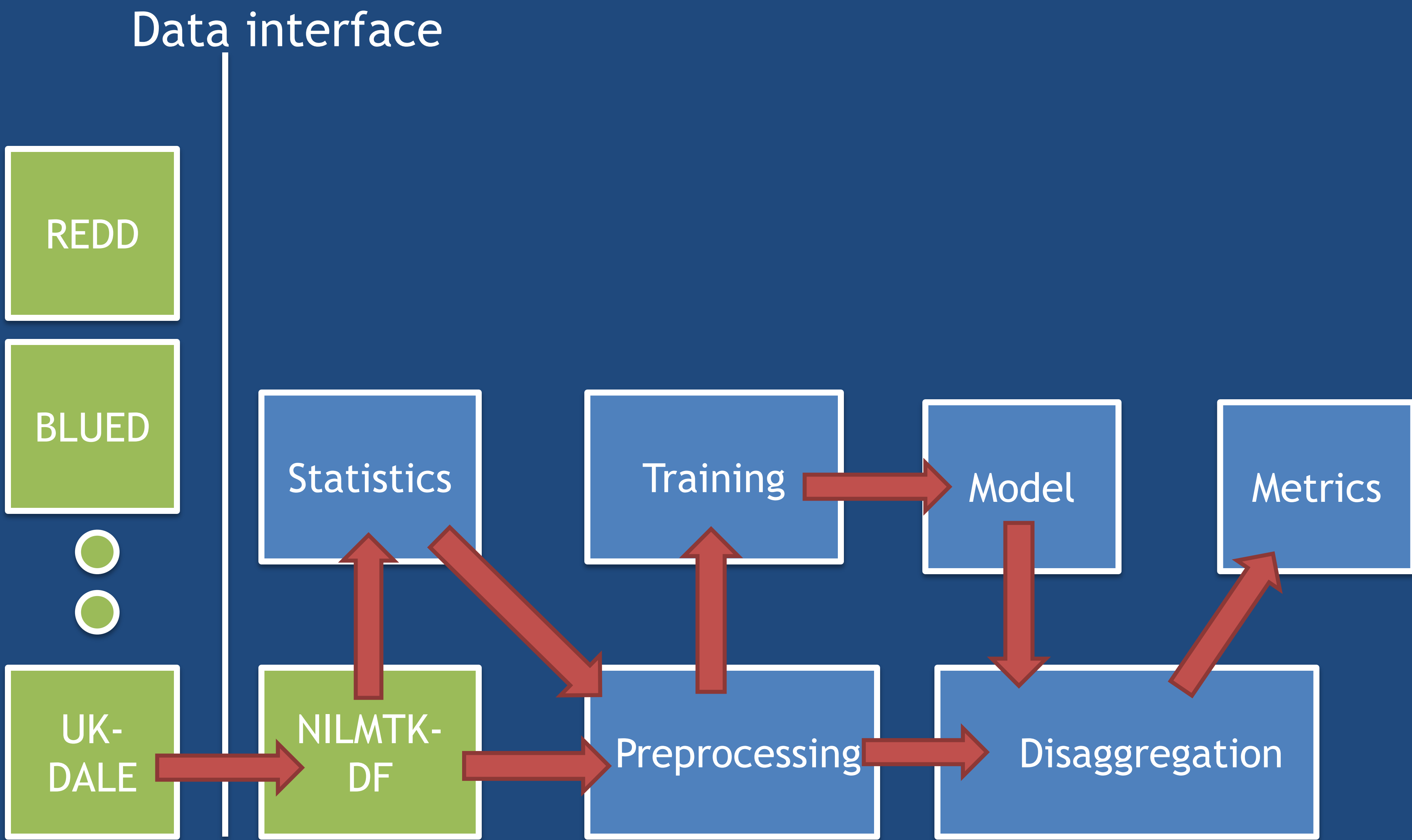
How does it do that?

Provides a pipeline from data sets to metrics to lower the entry barrier for researchers.

NILMTK pipeline



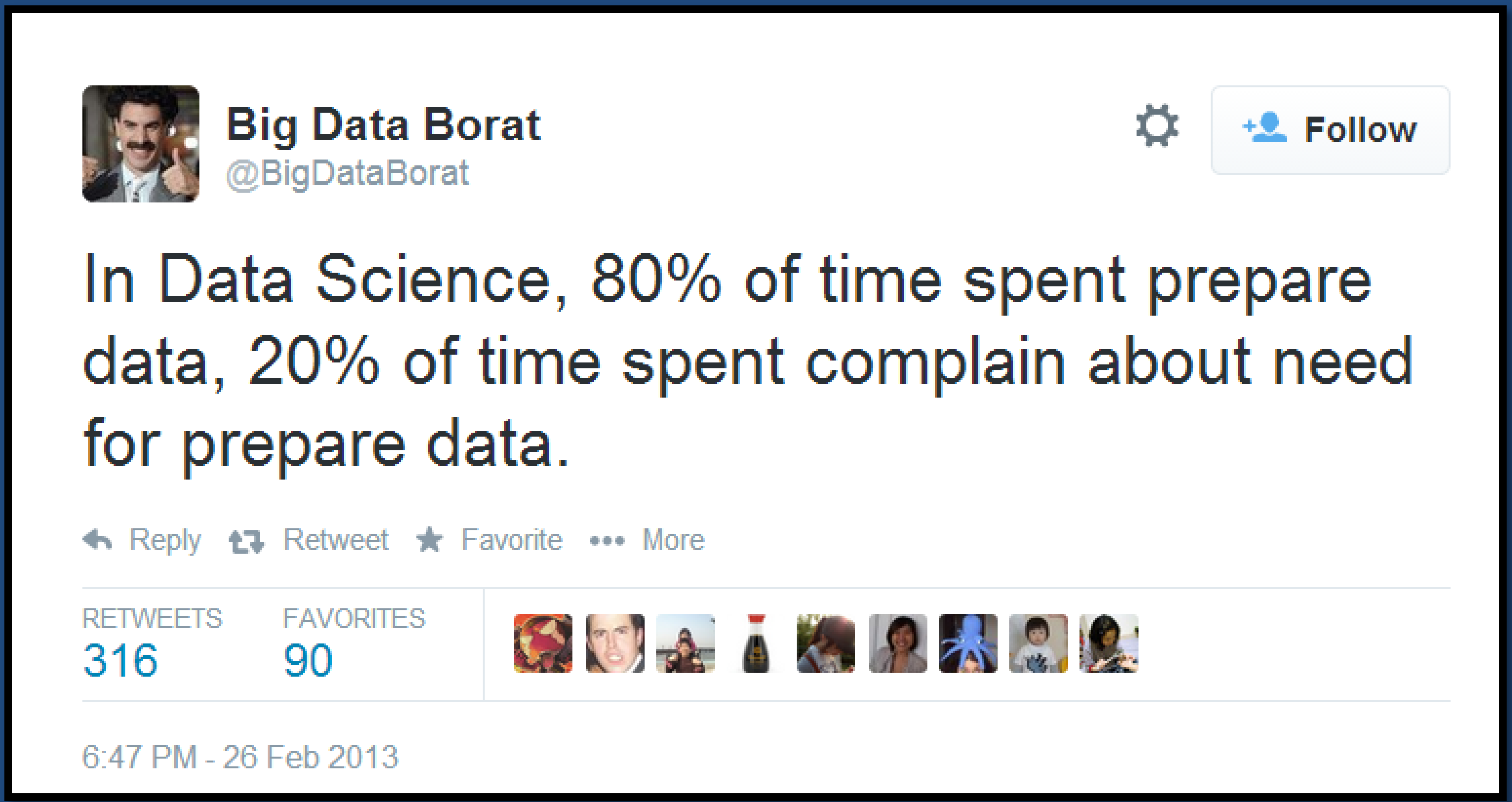
Data Format


Data interface
REDD
BLUED
UK-DALE
NILMTK-DF
Statistics
Training
Model
Metrics
Preprocessing
Disaggregation

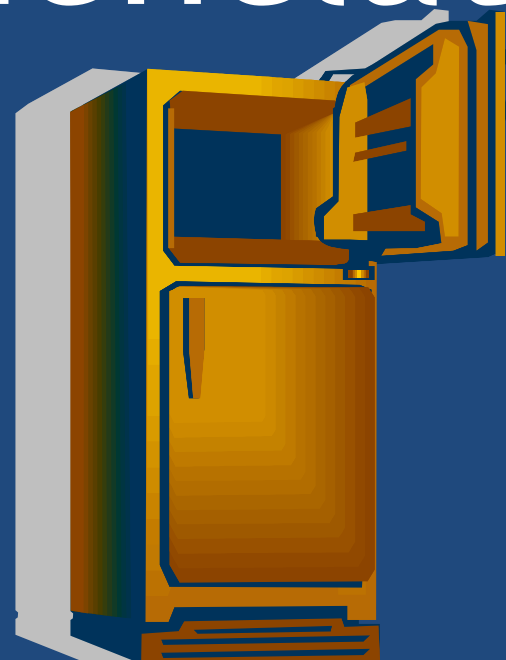
Data Format

- We propose NILMTK-DF: a common data format.
- Provide importers for 6 datasets: REDD, SMART*, Pecan street, iAWE, AMPds, UK-DALE
- Both flat file and efficient binary storage format

The fun of data!



Standardizing nomenclature

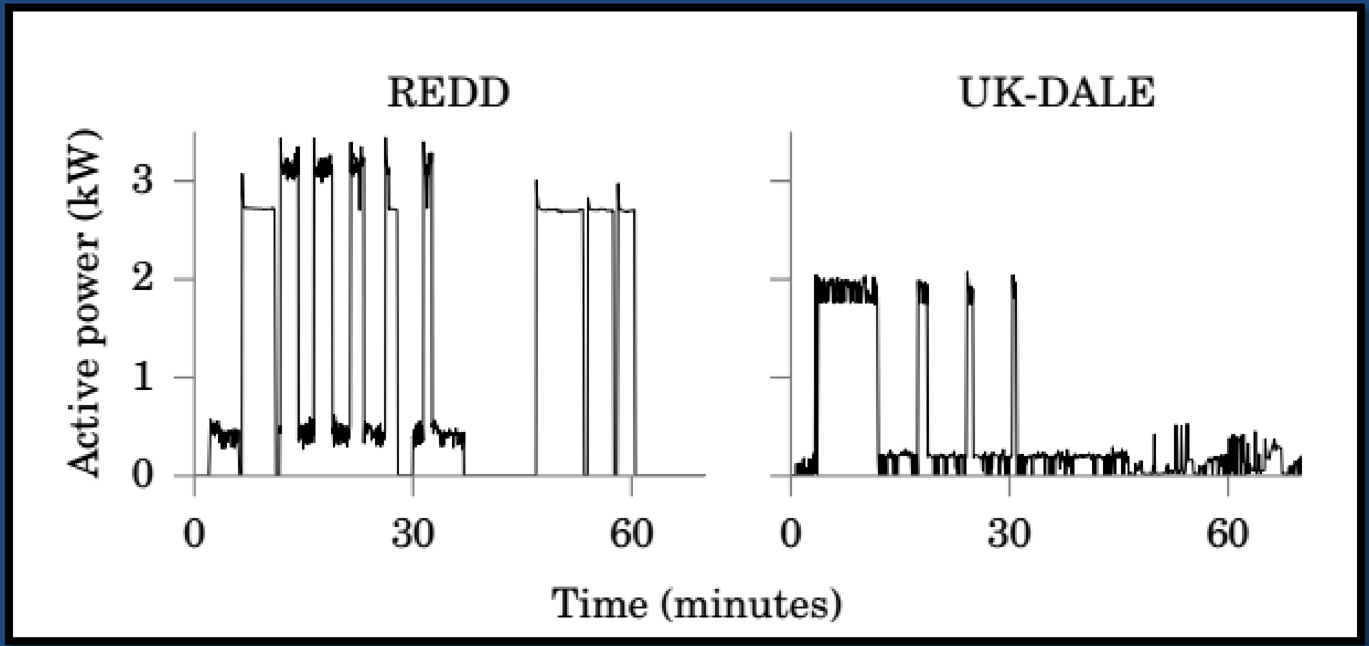


Metadata

- Geographic coordinates
- Type of appliance- hot, cold, dry?
- Metering hierarchy
- Parameters measured

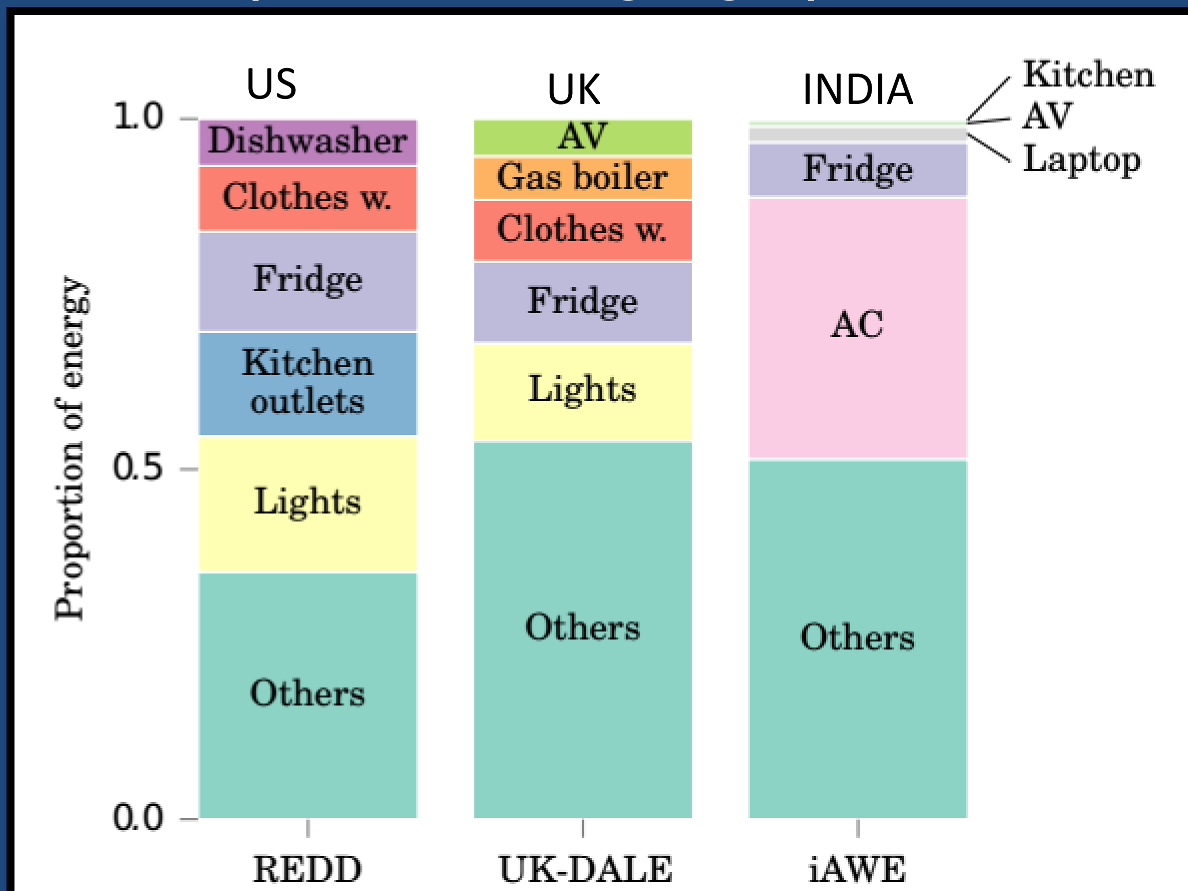
Standard nomenclature + Metadata + Datasets =

Comparing power draw of washing machines across US (REDD) and UK (UK-DALE)

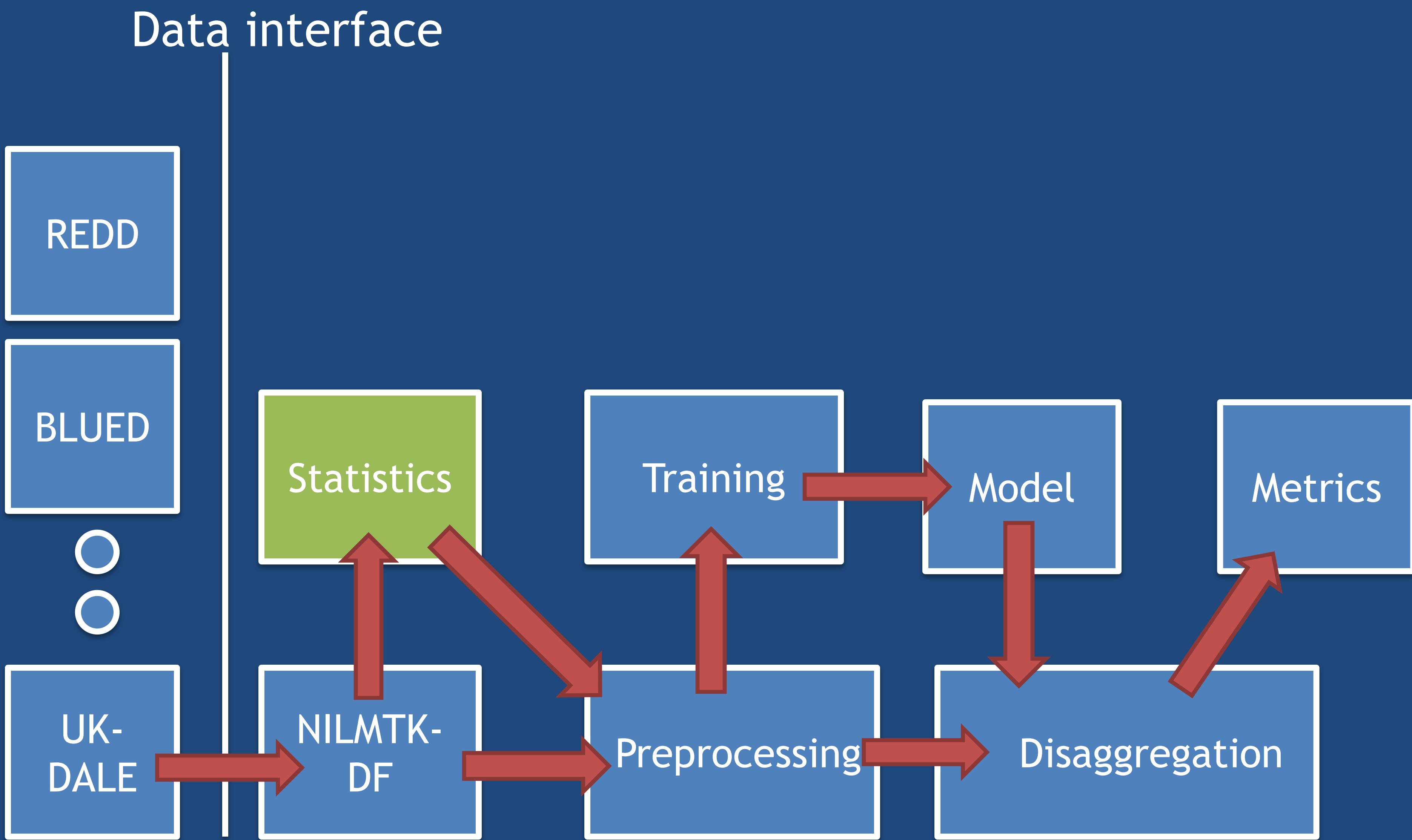


Standard nomenclature + Metadata + Datasets =

Top 5 appliance according to energy consumption across geographies

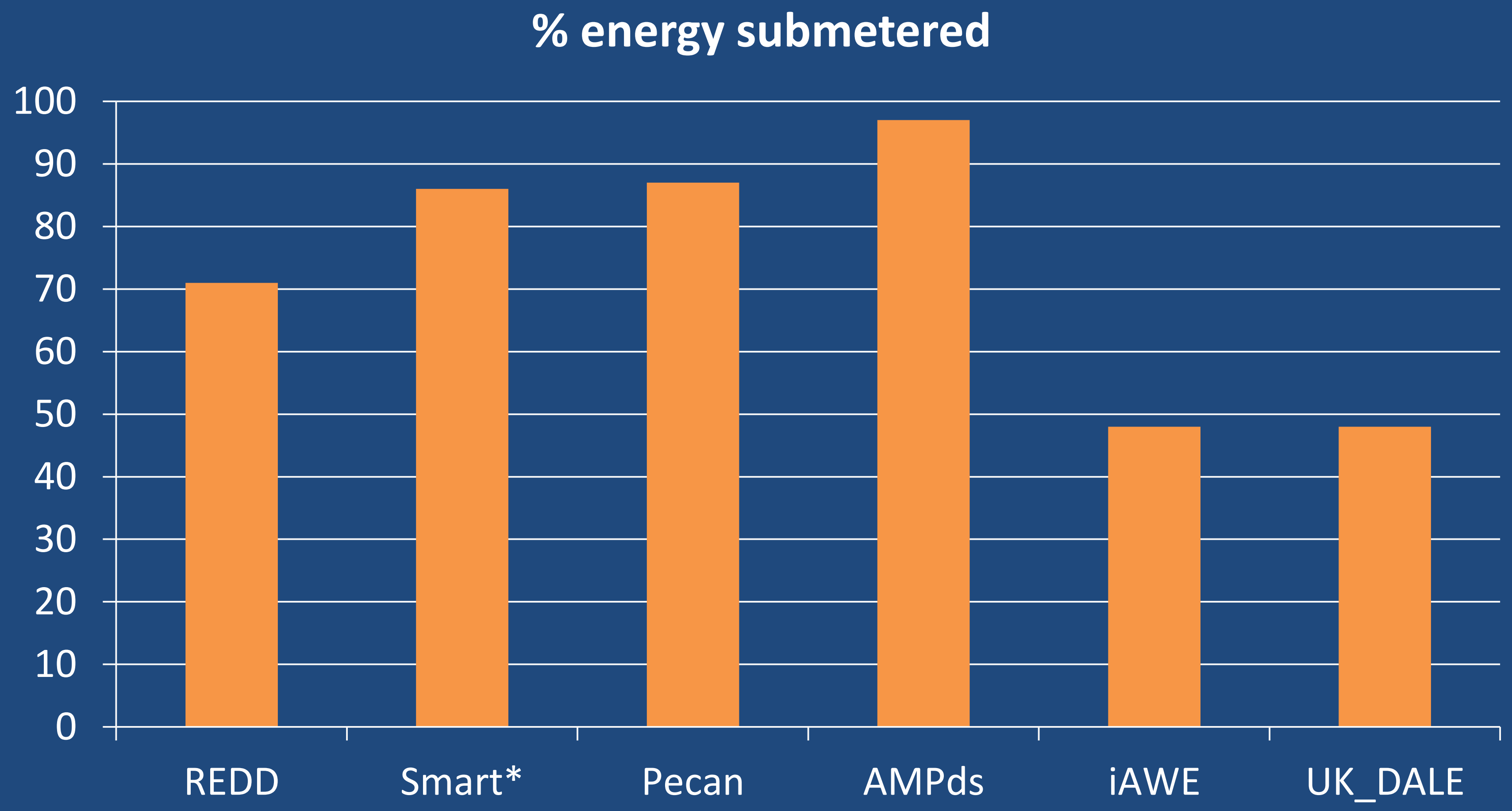


NILMTK pipeline



Statistics

- Energy submetered: Sum of energy of all appliance/Energy at mains level
- More energy submetered → More ground truth

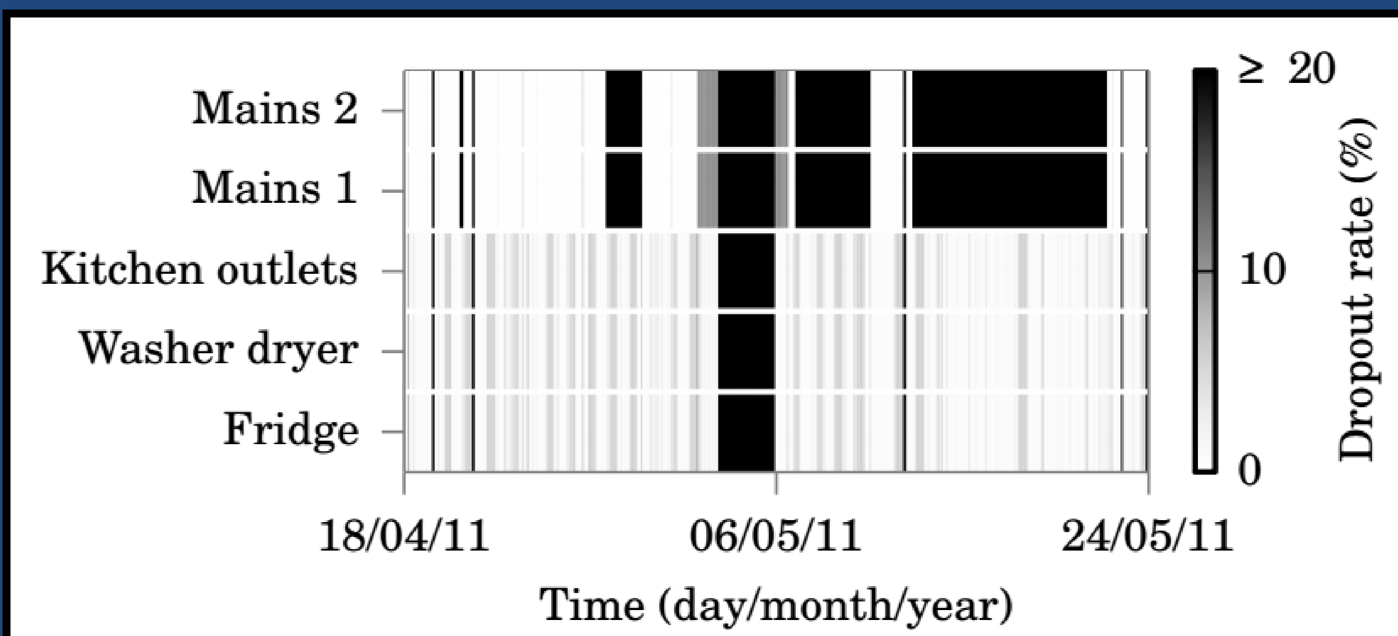


Statistics

- Appliance usage patterns
- Correlations with weather
- Appliance power demands

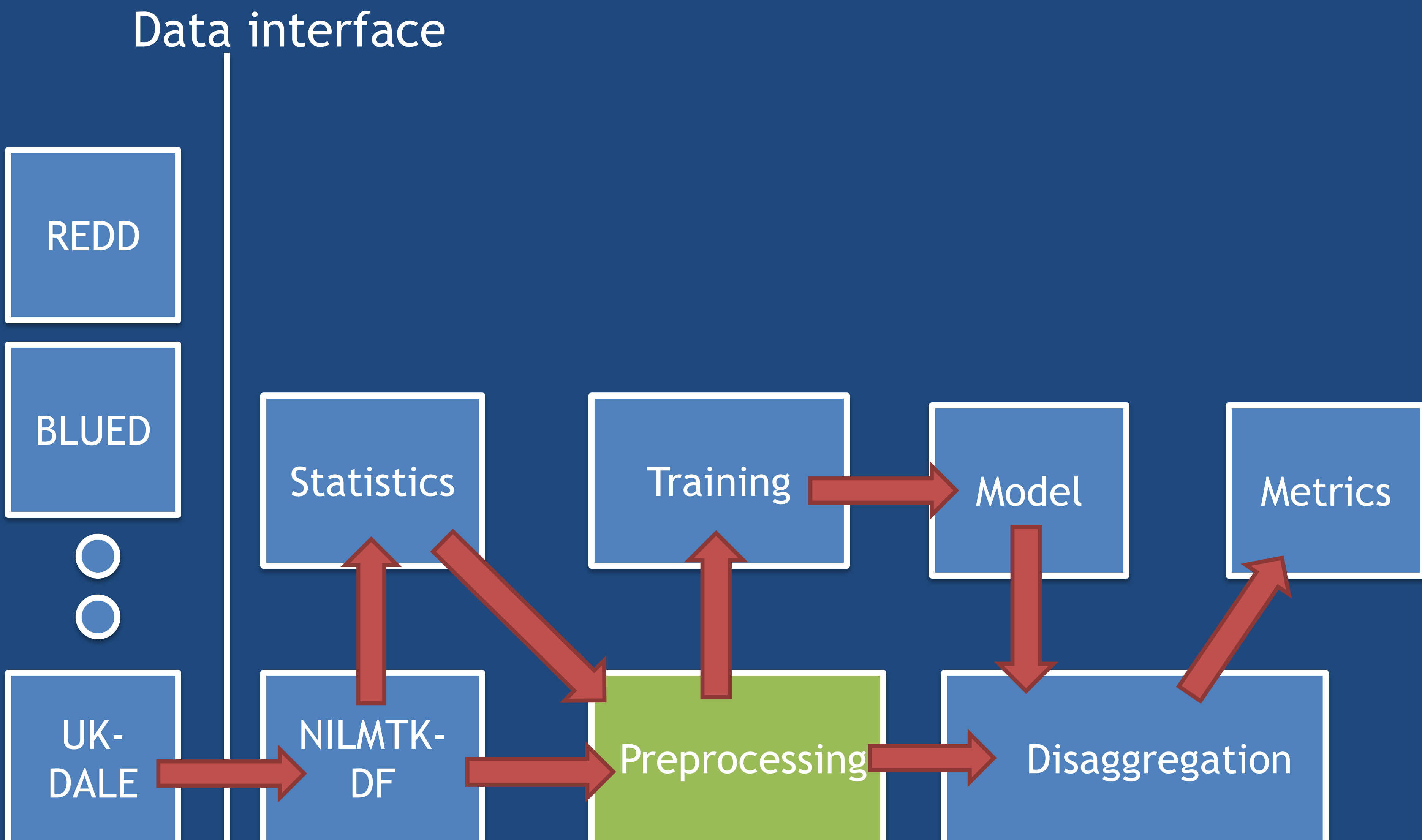
Diagnostics

- Every data set has problems ☹ NILMTK provides diagnostic functions for common problems.
- %Lost samples (per interval and whole), uptime



% lost samples in house 1 of REDD dataset

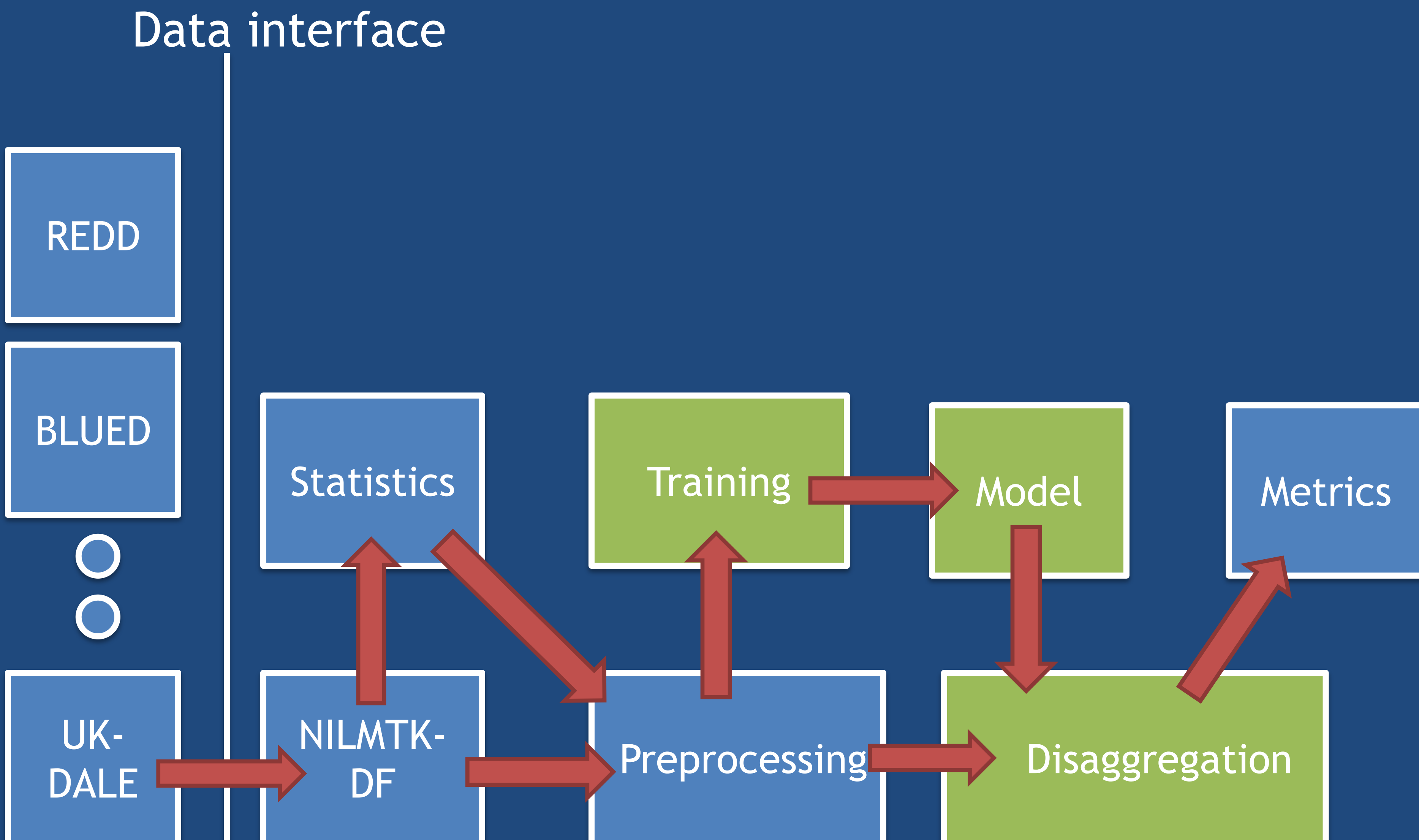
Preprocessing



Preprocessing

- Correct common problems (as per diagnosis).
- Other standard NILM preprocessors:
  - Interpolating, filtering implausible
  - Downsample to lower frequency
  - Select Top-k-appliances by energy consumption

Heart of NILMTK



Training

- NILMTK provides two benchmark algorithms
  - Combinatorial optimization (CO) [Proposed by Hart]
  - Factorial hidden Markov model (FHMM) [More recent, more complex]

Model

- Beyond the usual train and disaggregate, NILMTK allows importing and exporting learnt models
- Allows NILM to be deployed in “real world settings”
- Action speaks louder than words!! Demo follows!

Disaggregate!

- Quite a bit of work before we disaggregate
- We performed
  - CO and FHMM based disaggregation across first home of each dataset
  - Detailed disaggregation analysis across the home in iAWE (dataset from India)

Disaggregation across multiple datasets

- CO as good as FHMM across iAWE, UKPD, Pecan datasets
  - Space heating contributes 60% in Pecan and 35% in iAWE. Both approaches able to detect with fair ease

And I thought that CO was really outdated…

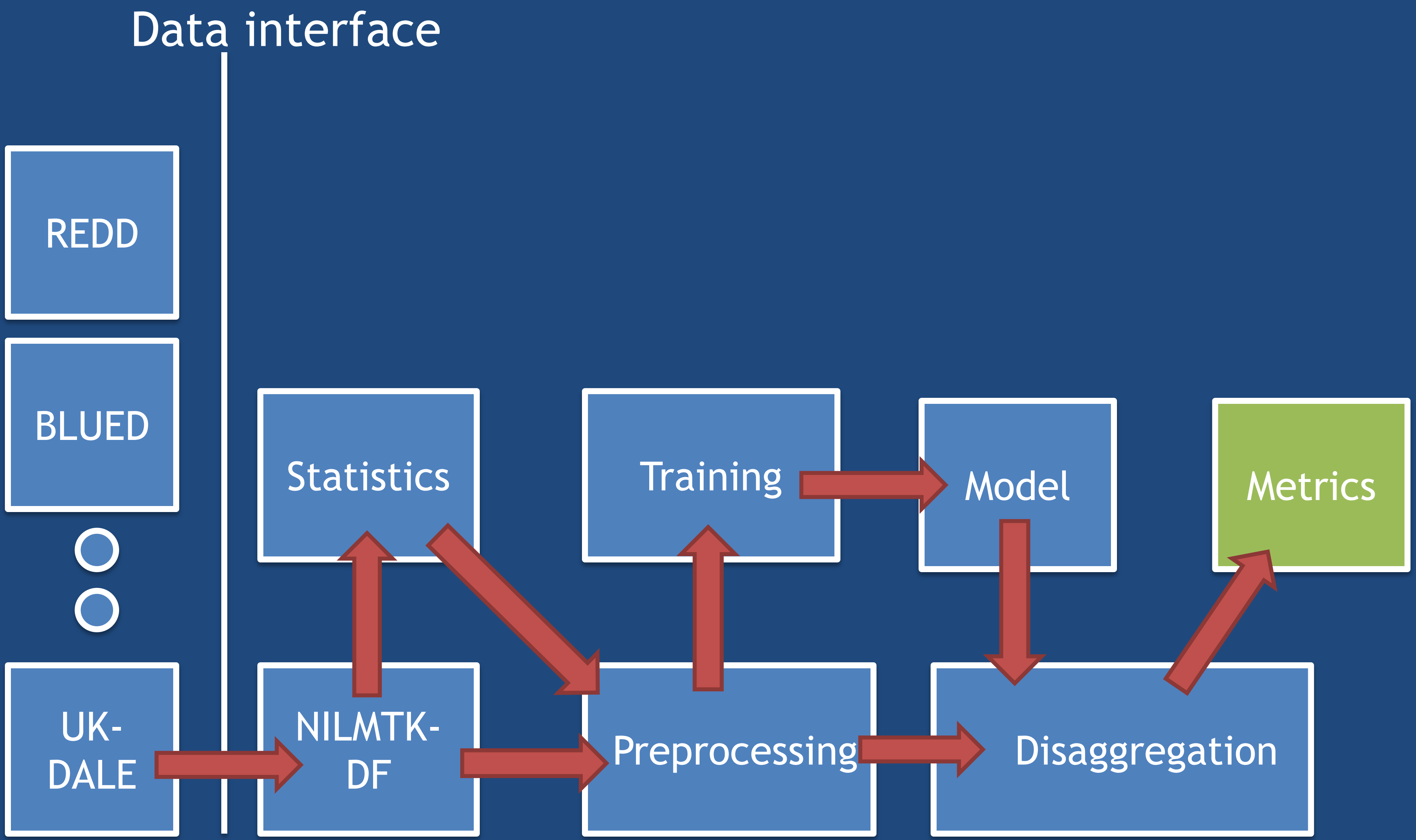
Disaggregation across multiple datasets

- FHMM outperforms CO across REDD, Smart*, AMPds
  - This is expected as FHMM models time variations.
- CO exponentially quicker than FHMM

Detailed disaggregation in iAWE dataset (India)

- CO and FHMM perform similar
- Appliances such as air conditioners way easier to disaggregate
- Complex appliances (laptops and washing machines) – not so good ☹

NILMTK pipeline



Metrics

- NILMTK provides:
  - General machine learning metrics
    - Precision, Recall, F-score
  - Specialized metrics for NILM
    - Error in total energy assigned, RMS error in assigned power,..
  - Both event based and total power based NILM metrics.

Demo time!!

Conclusions

Three core challenges in NILM research

1. Hard to address generality
2. Lack of comparison against same benchmarks
3. Inconsistent disaggregation performance metrics

How NILMTK addresses these challenges

1. Standard input and output formats (Addresses #1)
2. Parsers for 6 NILM data sets (Addresses #1, #2)
3. Two benchmark NILM algorithms (Addresses #1, #2)
4. Statistics, diagnostics and preprocessing (Addresses #1, #2)
5. Metrics for different NILM use cases (Addresses #1)

Backup

Combinatorial optimization

- Seeks to find the optimal combination of appliances' power draw to minimize residual energy.
- Similar to subset-sum problem and thus NP-complete ☹
- Power draw is not related in time

Combinatorial optimization

Appliance	Off power	On power
Air conditioner (AC)	0	2000
Refrigerator	0	200

If total power observed = 210 → AC is OFF and Refrigerator is ON

Combinatorial optimization

Appliance	Off power	On power
Air conditioner (AC)	0	2000
Refrigerator	0	200

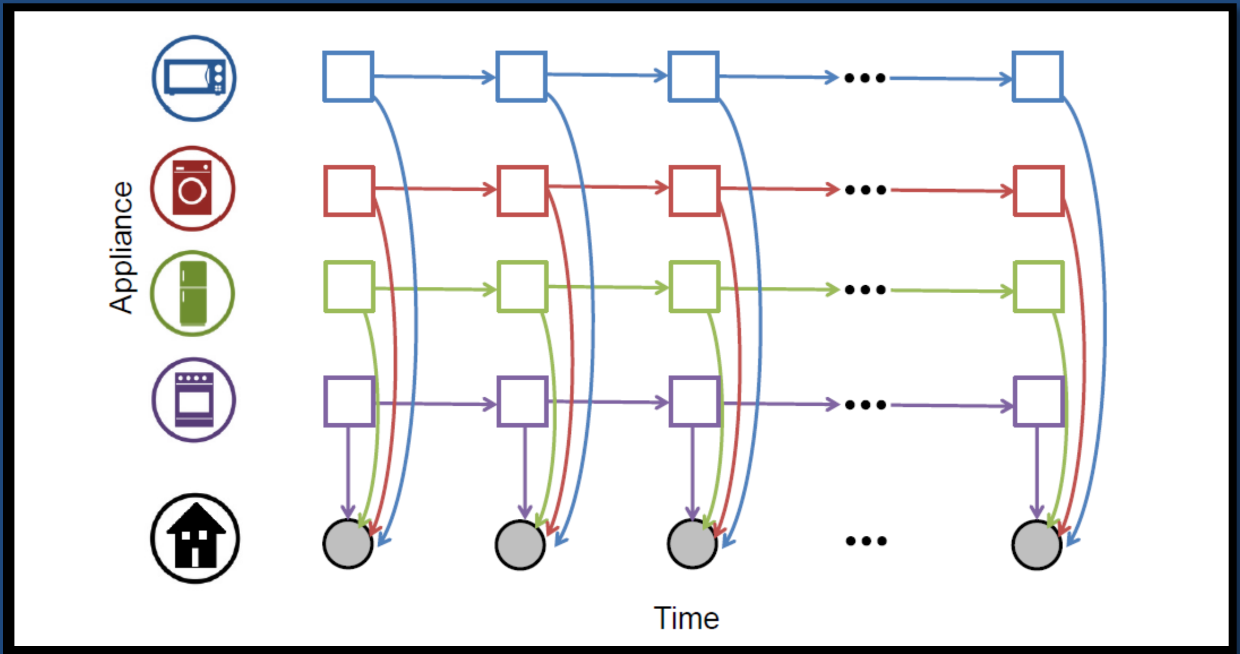
If total power observed = 2000 → AC is ON and Refrigerator is OFF

Combinatorial optimization

Appliance	Off power	On power
Air conditioner (AC)	0	2000
Refrigerator	0	200

If total power observed = 2230 → AC is ON and Refrigerator is ON

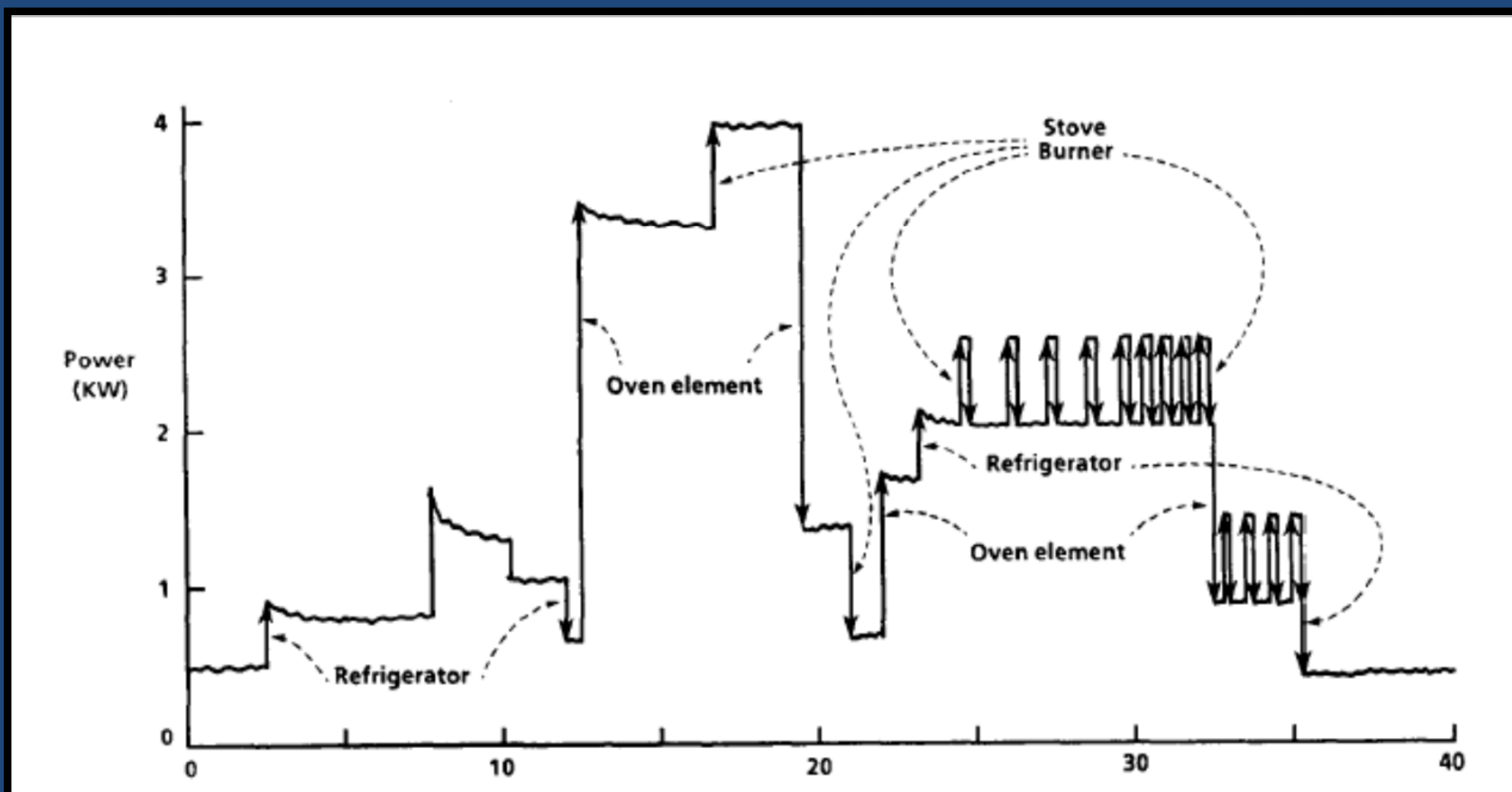
FHMM



- Each appliance modeled as HMM
  - Power draw related in time→ If TV is on right now, likely to be on next second.
- Exact inference scales worse than CO

A bit of history

Seminal work on NILM done at MIT dates back to early 1980s – A good 6-7 years before I was born!



Field progress