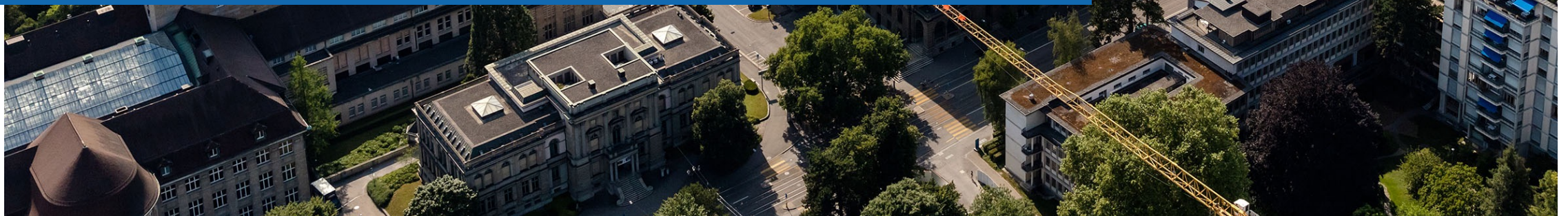


# Smart Meter Data Analytics:

Practical Use-Cases and Best Practices of Machine Learning Applications for Energy Data in the Residential Sector

**Tobias Bruder Müller & Markus Kreft**

11th International Conference on Learning Representations (ICLR 2023)



# The Bits to Energy Lab

**ETH**

Eidgenössische Technische Hochschule Zürich  
Swiss Federal Institute of Technology Zurich



Otto-Friedrich-Universität Bamberg

**FAU**

FRIEDRICH-ALEXANDER  
UNIVERSITÄT  
ERLANGEN-NÜRNBERG



University of St.Gallen

- **Team:** ~20 researchers with different professional backgrounds at 4 institutions (Germany & Switzerland)
- **Goal:** Leverage technology to foster sustainability
- **Focus:** Machine learning, energy applications, and digital behavioral interventions
- **Web:** [www.bitstoenergy.com](http://www.bitstoenergy.com) / [www.im.ethz.ch](http://www.im.ethz.ch)

IoT/  
Sensor Data



Machine  
Learning



Behavioral  
Interventions



Digital  
Transformation



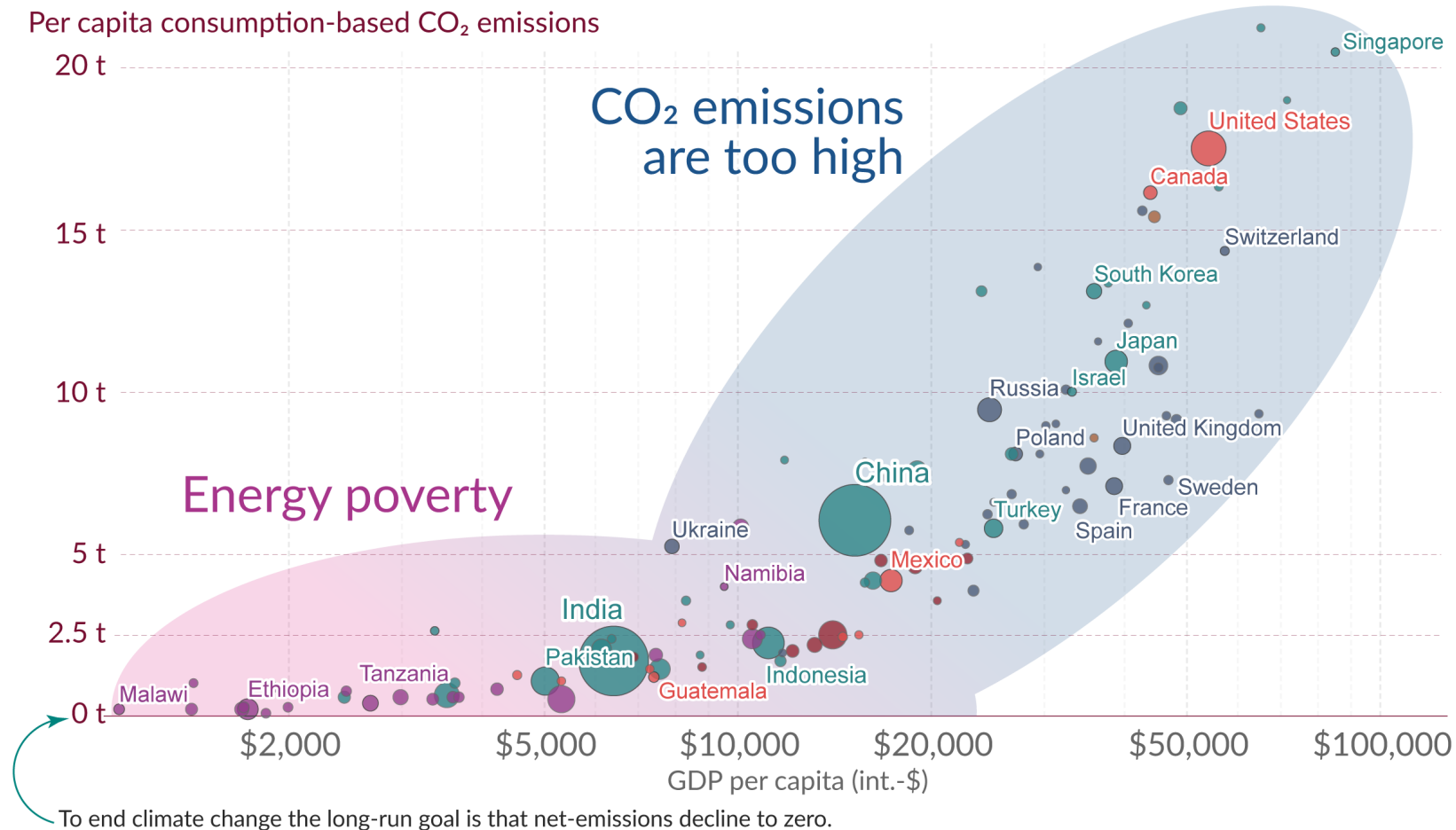


# United Nations - Sustainable Development Goals (SDGs)



Source: United Nations

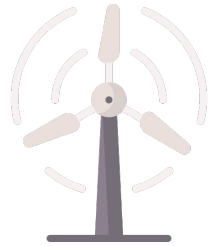
# Reality: CO2 Emissions per Capita vs. GDP per Capita



Source: Our World in Data (2020): <https://ourworldindata.org/worlds-energy-problem> - Last accessed: 2023 April 05

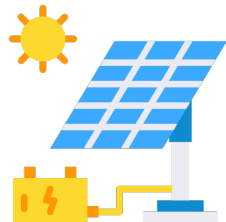
# Energy System - Electrification & Transformation (1)

## Generation



**+ 55 %**

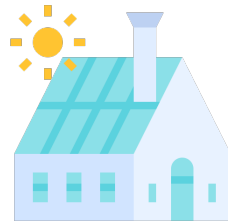
Wind electricity generation (2021 vs. 2020) [1]



**+ 22 %**

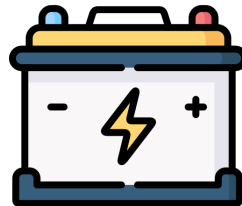
Solar PV generation (2021 vs. 2020) [2]

## Distribution



**70 %**

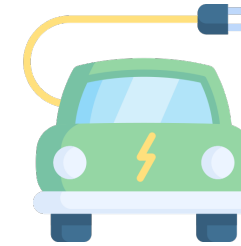
Of power transformers & transmission lines in the U.S. at least 25 years old (2015)



**+ 88 %**

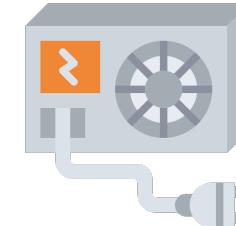
Global annual grid-scale battery storage additions (2021 vs. 2020) [4]

## Consumption



**13 %**

Of new cars sold are electric (2022) [5]



**+ 11 %**

Global annual growth of heat pump sales (2021 & 2022) [6]

[1] IEA (2022) <https://www.iea.org/reports/wind-electricity>, Last accessed: 2023 April 05

[2] IEA (2000-2021) <https://www.iea.org/data-and-statistics/charts/evolution-of-annual-solar-pv-installations-and-share-by-segmentation-2000-2021>, Last accessed: 2023 April 05

[3] IEA (2022) <https://www.iea.org/reports/unlocking-the-potential-of-distributed-energy-resources>, Last accessed: 2023 April 05

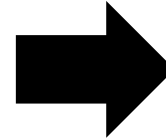
[4] IEA (2016-2021) <https://www.iea.org/data-and-statistics/charts/annual-grid-scale-battery-storage-additions-2016-2021>, Last accessed: 2023 April 05

[5] IEA (2022) <https://www.iea.org/reports/electric-vehicles>, Last accessed: 2023 April 05

[6] IEA (2021-2022) <https://www.iea.org/data-and-statistics/charts/annual-growth-in-sales-of-heat-pumps-in-buildings-worldwide-and-in-selected-markets-2021-and-2022>, Last accessed: 2023 April 05

# Energy System - Electrification & Transformation (2)

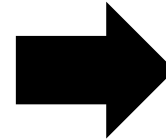
- Population growth
- Increasing access to electricity
- Electrification in all sectors



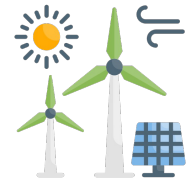
**Higher electricity demand**



- Increasing share of renewables
- Higher degree of decentralization



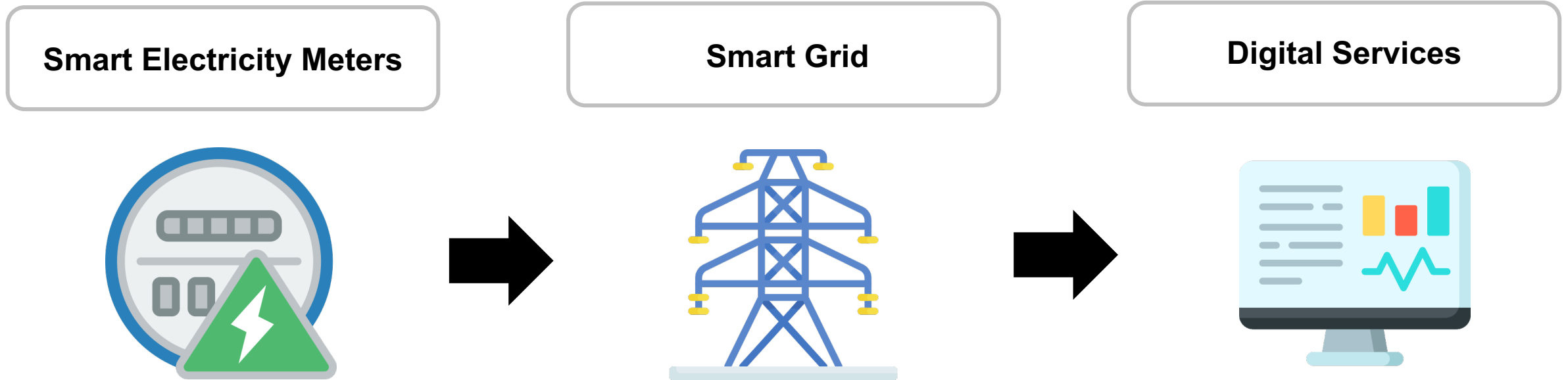
**Larger fluctuations in electricity production**



**Need for ....**  
**...smart control, energy storage, energy efficiency, etc...**



# Part of Solution: The Advanced Metering Infrastructure (AMI)



- High penetration, e.g.:
  - U.S.: 50% (2016) [1]
  - China, India, Japan, South Korea: 69% (2019) [2]
  - E.U.: 80% (2020-2025) [3]

- Monitoring activities on the grid
- Decentralized control
- Demand response programs
- Planning & forecasting
- Dynamic tariff design
- etc.

- Energy efficiency insights
- Smart buildings & retrofits
- Electric vehicles charging
- Heat pump operation
- etc.

[1] U.S. Energy Information Administration (2027) <https://www.eia.gov/todayinenergy/detail.php?id=34012>, Last accessed: 2023 April 06

[2] Smart Energy International (2021) <https://www.smart-energy.com/industry-sectors/smart-meters/smart-electricity-meters-rollout-in-china-india-japan-and-south-korea/>, Last accessed: 2023 April 06

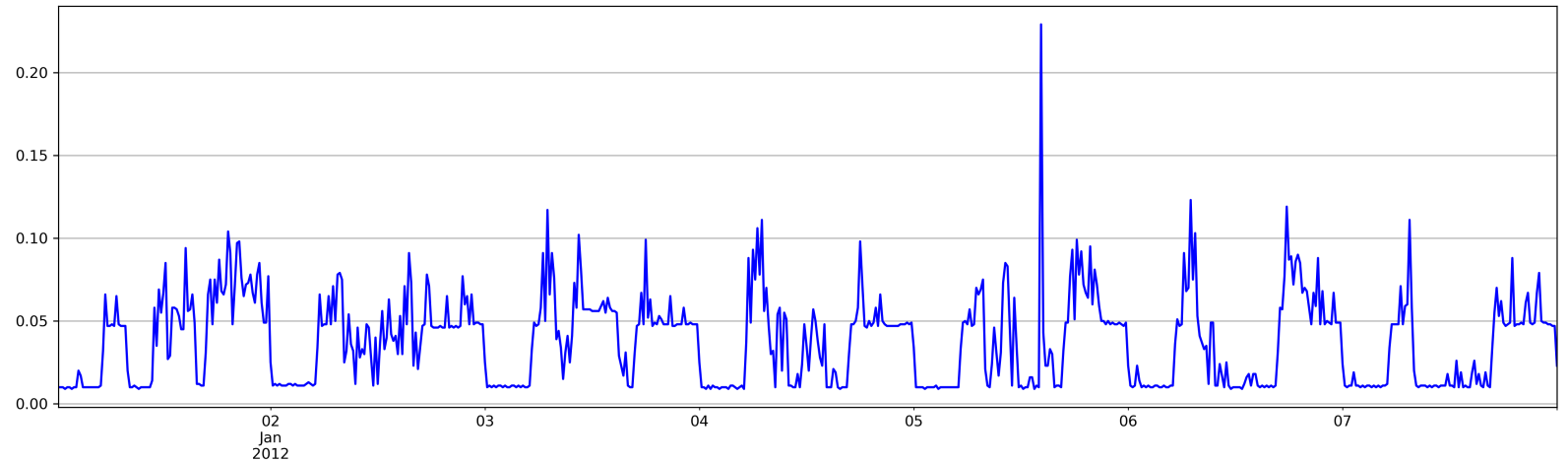
[3] Efkarpidis, N., Geidl, M., Wache, H., Peter, M., & Adam, M. (2022). Smart Metering Applications. In Smart Metering Applications: Main Concepts and Business Models (pp. 13-124). Cham: Springer International Publishing.

# Example: Smart Meter Data (SMD)



**Smart Meter Data:**  
Time series of power or energy measurements at different resolutions (~1s to 1h)

## Electrical Energy (kWh)



Time (15 min resolution)



# The Goal of our Tutorial

## Insights Into Energy Domain

- Current trends & use-cases in energy space
- Typical & atypical energy use in residential buildings
- Value of smart meter data & smart grids



## Educating About Data Analysis

- Practical, self-paced learning in a “hands-on” manner
- Applying best practices, ML & data mining to real-world energy data
- Provide starting point for analyzing own energy data



## Creating Awareness

- For energy efficiency, own energy use and climate change
- For novel data sets, methods, and applications in ML community
- For necessity for exchange between research communities



```
1.6 Data availability
1.7 Contextual information and domain knowledge
1.8. Making use of additional data sources
2. Preparations for this tutorial
2.1 Importing Packages
2.2 Importing the data
3. Best practices for visualizing smart meter data
3.1 Time-series visualizations of energy data
3.2 Visualizing distributions of energy consumption
3.3 Multi-dimensional visualizations
3.4 Annotating visualizations with additional context information
3.5 Displaying aggregated demand
4. Pre-processing smart meter data
4.1 Combining smart meter data with temperature data
4.2 Add additional information for filtering timestamps
4.3 Normalization methods
4.4 Interpolation methods and downsampling
4.5 Outlier detection with Hampel filter
4.6 Simple baseload estimation
4.7 Enhancing small activities
4.8 Detecting switching activities
4.9 Sliding window approaches
4.10 Feature extraction
4.11 Detecting peaks in distribution
4.12 Frequency-based methods for low-resolution data
5. Non-Intrusive Load Monitoring (NILM) / Load Disaggregation
5.1 Using classification algorithms to detect appliance installations
5.2 Applying simple deep learning models for NILM
5.3 Applying Hidden Markov Models for NILM
```

```
plt.figure(figsize=(12,8))
plt.plot(mo_temp, mo_energy, 'o', color='blue', alpha=0.3)
plt.plot(e_hat, df_energy.signature['usage_hat'], '-', color='orange')
plt.ylabel('Normalized Energy Consumption', fontsize=14, fontweight='bold')
plt.xlabel('Outside Air temperature [°C]', fontsize=14, fontweight='bold')
plt.title('Step 2: Fit Non-Linear Model (Score: {:.3})'.format(np.round(score, 2)), fontsize=16)
plt.show()
```

Step 1: Scatter Plot and Outlier Detection

Step 2: Fit Non-Linear Model (Score: -1.5)

8.2 Extraction of average day profiles

In many customer segmentation applications, an **average daily profile** is calculated per customer. Such a profile is given by the **average consumption at each time of day**. The average daily profiles can serve as input to a classification or clustering algorithm. Alternatively, such a profile can also be useful for calculating how much an individual household's energy consumption fluctuates around the average profile, i.e., how regular the patterns are from day to day. More sophisticated methods could further distinguish between days of the week or seasons, but below we will look at a very simple example that extracts the average daily profile of all summer days available in our smart meter data.

```
[ ] # DEFINE HELPER METHODS
# calculate_average_day_profile(df):
def calculate_average_day_profile(df):
    """
    Calculates an average day profile (i.e. the mean consumption of every time of the day).
    Args:
        df: data frame to be analyzed
    Returns:
        data frame with average day profile
    """
    #_copy_data frame such that it is not affected by changes
```

# Topics Covered (1)

## Introduction to smart meter data

1. What is smart meter data?
2. Relevance for tackling climate change
3. Chances and limitations
4. Power vs. energy measurements
5. Data resolution
6. Data availability
7. Contextual information and domain knowledge
8. Making use of additional data sources



## Best practices for visualizing smart meter data

1. Time-series visualizations of energy data
2. Visualizing distributions of energy consumption
3. Multi-dimensional visualizations
4. Annotating visualizations with additional context information
5. Displaying aggregated demand



## Pre-processing smart meter data

1. Combining smart meter data with temperature data
2. Add additional information for filtering timestamps
3. Normalization methods
4. Interpolation methods and downsampling
5. Outlier detection with Hampel filter
6. Simple baseload estimation
7. Enhancing small activities
8. Detecting switching activities
9. Sliding window approaches
10. Feature extraction
11. Detecting peaks in distribution
12. Frequency-based methods for low-resolution data

# Topics Covered (2)

## Non-Intrusive Load Monitoring (NILM)

1. Using classification algorithms to detect appliance installations
2. Applying simple deep learning models for NILM
3. Applying Hidden Markov Models for NILM
4. Correctly evaluating NILM approaches
5. Rule-based heuristics for pattern isolations



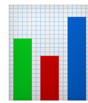
## Anomaly Detection

1. Introduction to online vs. offline change point detection
2. Finding state changes with offline change point detection
3. Finding anomalies with sliding window
4. Applying Symbolic Aggregate ApproXimation
5. Finding discords and motifs



## Flexibility Estimation

1. Estimating load shifting potential of disaggregated appliances



## Load Forecasting

1. Brief and short introduction



## Customer segmentation

1. Extraction of energy signatures through regression
2. Extraction of average day profiles
3. Applying clustering algorithms

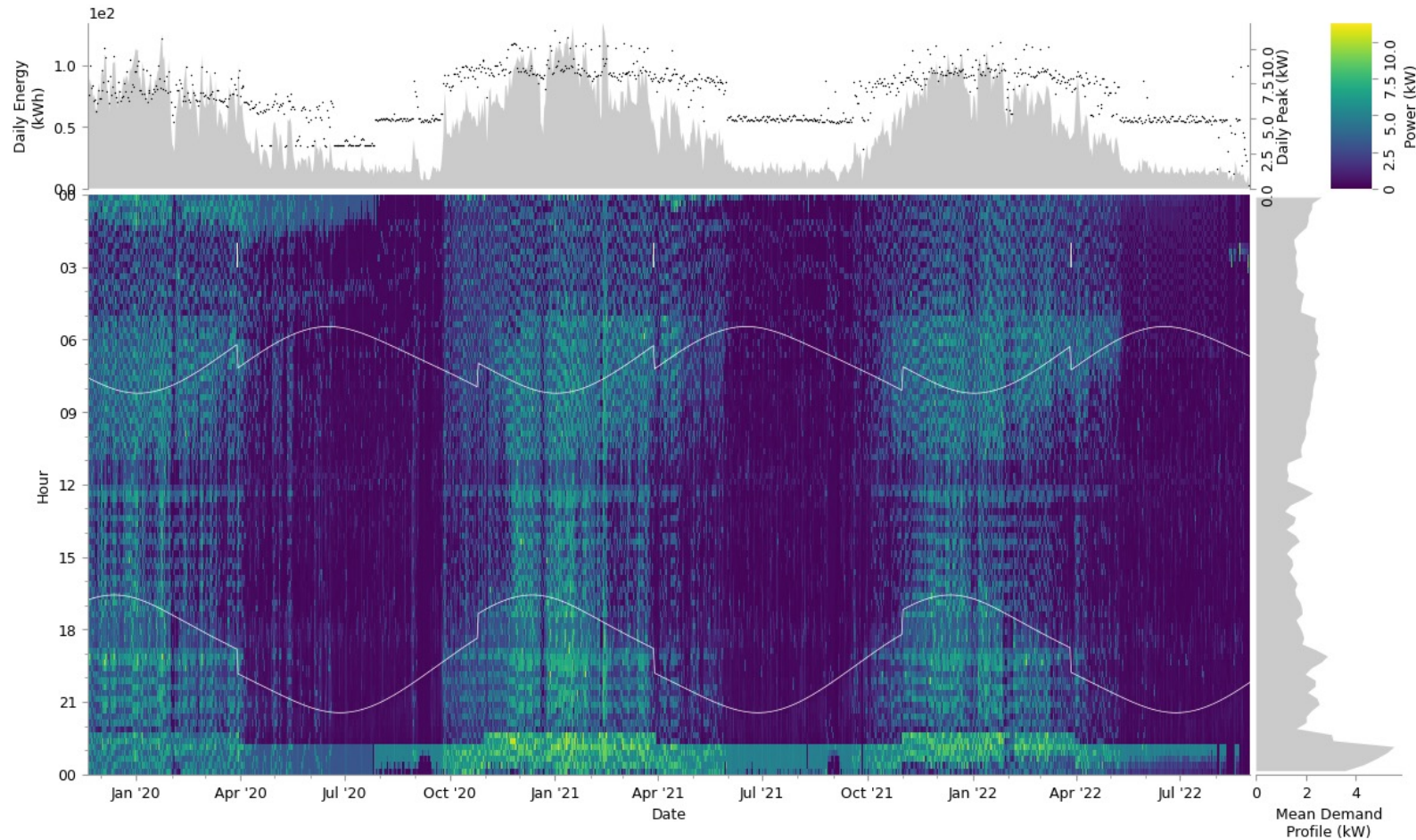


## Additional Resources

1. Data Sets
2. Software Packages
3. Other



# Example: Visualization of Energy Consumption



**Visualization as Heat Map:**  
Observe regular energy  
consumption patterns over time



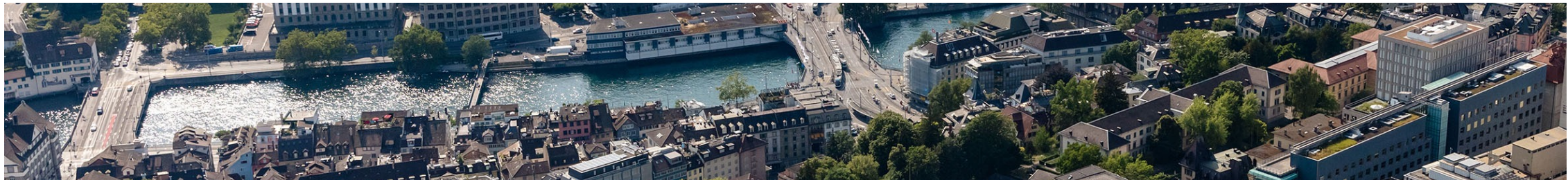


## **Tobias Bruder Müller & Markus Kreft**

ETH Zürich  
Bits to Energy Lab @ Chair of Information Management  
Department of Management, Technology, and Economics

Weinbergstr. 56/58  
8092 Zurich  
Switzerland

E-Mail: [tbrudermuell@ethz.ch](mailto:tbrudermuell@ethz.ch) & [mkreft@ethz.ch](mailto:mkreft@ethz.ch)  
Web: [www.bitstoenergy.com](http://www.bitstoenergy.com) / [www.im.ethz.ch](http://www.im.ethz.ch)



# Smart Meter Data Analytics:

Practical Use-Cases and Best Practices of Machine Learning Applications for Energy Data in the Residential Sector

**Tobias Bruder Müller & Markus Kreft**

11th International Conference on Learning Representations (ICLR 2023)

