

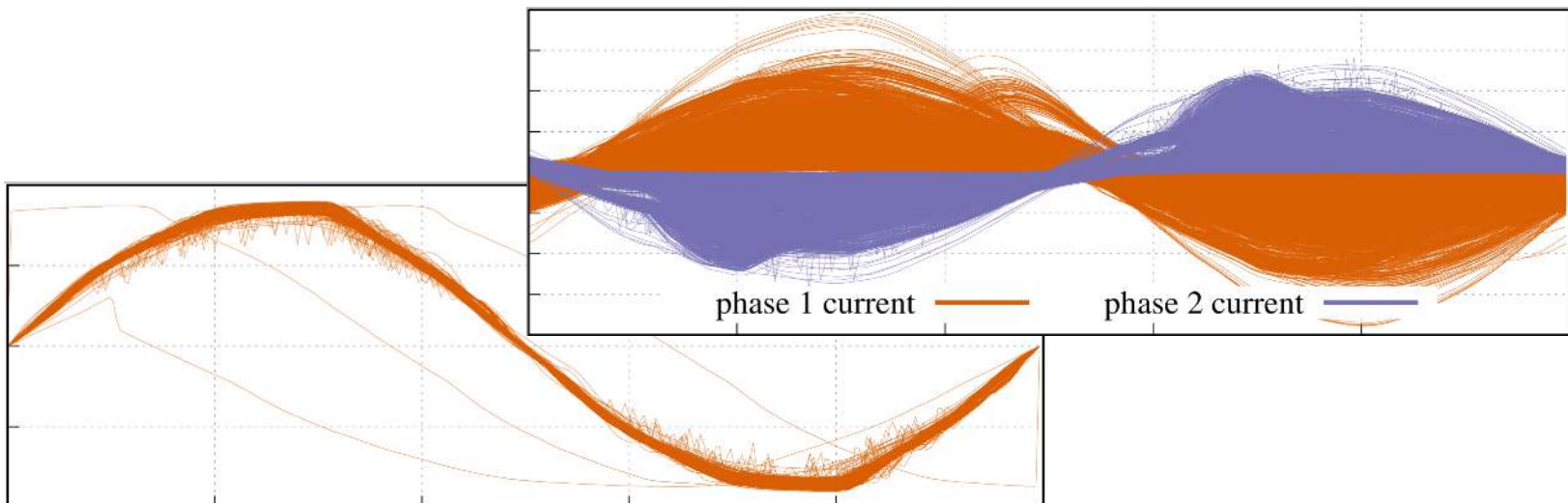


# Signalverarbeitung und maschinelles Lernen als Werkzeuge zur Verbesserung der Energieeffizienz

Dr.-Ing. Andreas Reinhardt

Leiter der Abteilung für Energieinformatik

Institut für Informatik, TU Clausthal



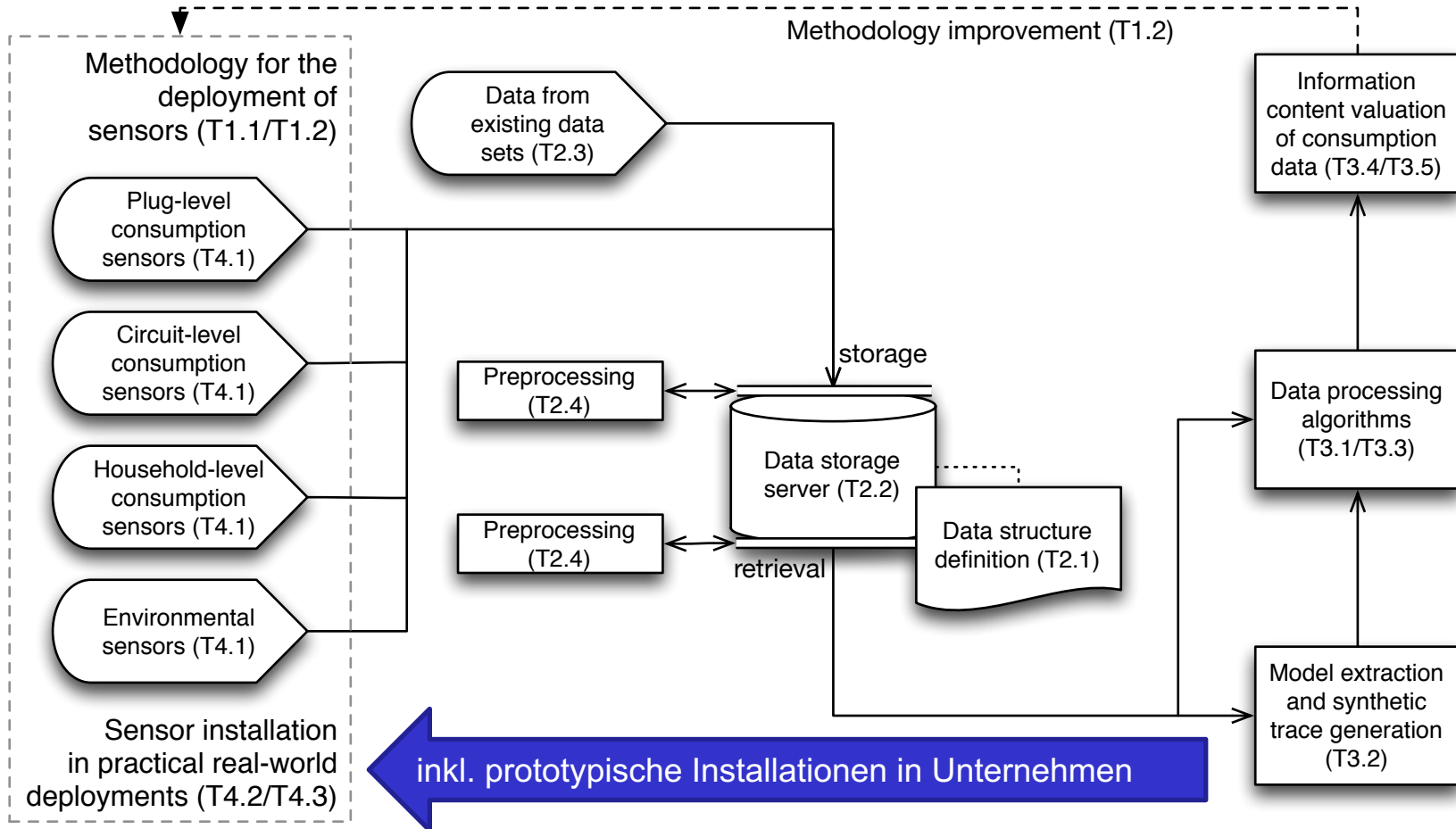
## Abteilung für Energieinformatik

- Institut für Informatik, TU Clausthal
  - Campus Tannenhöhe, Standort Clausthal-Zellerfeld
- Forschung und Entwicklung an Systemen zur Erfassung und Verarbeitung von Energieinformationen
  - Hochaufgelöste Datenerfassung aus multimodalen Quellen
  - Übertragung und Fusion der Messwerte (ggf. Vorverarbeitung)
  - Datenanalyse durch Signalverarbeitung und künstliche Intelligenz
- Förderung der Forschungsaktivitäten
  - Projekt SECoM: “A Systematic Energy Information Collection Methodology for Improved Energy Analytics”
  - Deutsche Forschungsgemeinschaft (RE 3857/2-1) und Simulationswissenschaftliches Zentrum SWZ

Gefördert durch  
**DFG** Deutsche  
Forschungsgemeinschaft

**SWZ** Clausthal - Göttingen  
Simulationswissenschaftliches Zentrum

## Projektüberblick SECoM



# Messaufbau zur multimodalen Datenerfassung

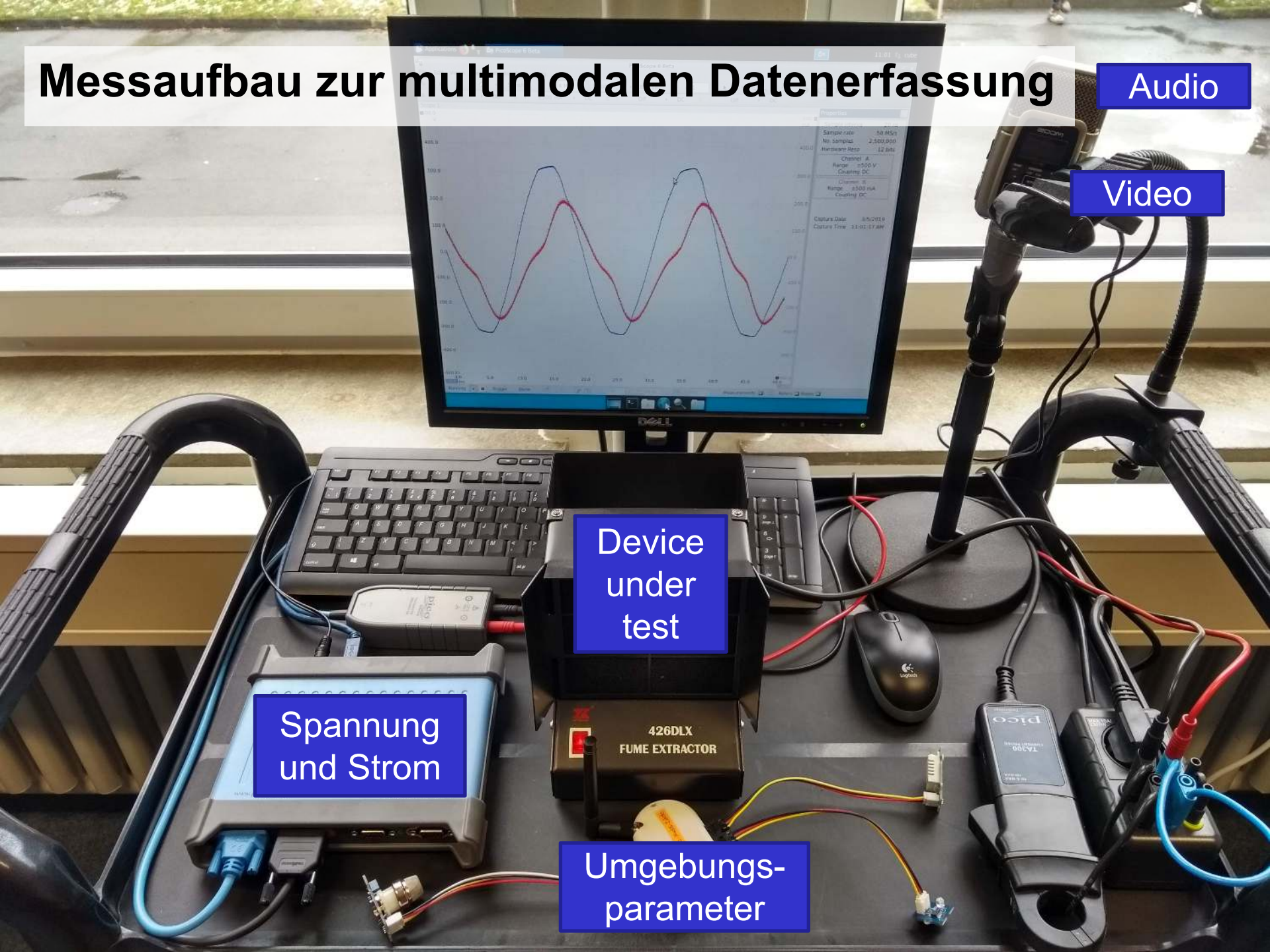
Audio

Video

Device  
under  
test

Spannung  
und Strom

Umgebungs-  
parameter





Aber nun zurück  
zum Thema

ZIEHEN SIE  
IM VORÜBERGEHEN  
#200 GEHALT EIN.

**LOS**



$$\text{Energieeffizienz} = \frac{\text{Ertrag}}{\text{Energieeinsatz}} \quad (\text{vgl. 2012/27/EU})$$

"Verbesserung der Energieeffizienz" → wirtschaftlicherer Einsatz elektrischer Energie



Bildquelle: <https://forsythesadvisors.com.au>

Hierzu ist jedoch zunächst die Kenntnis (besser noch: ein tiefgehendes Verständnis) des Energieeinsatzes erforderlich

# Wege zur Erfassung elektrischer Energieverbräuche



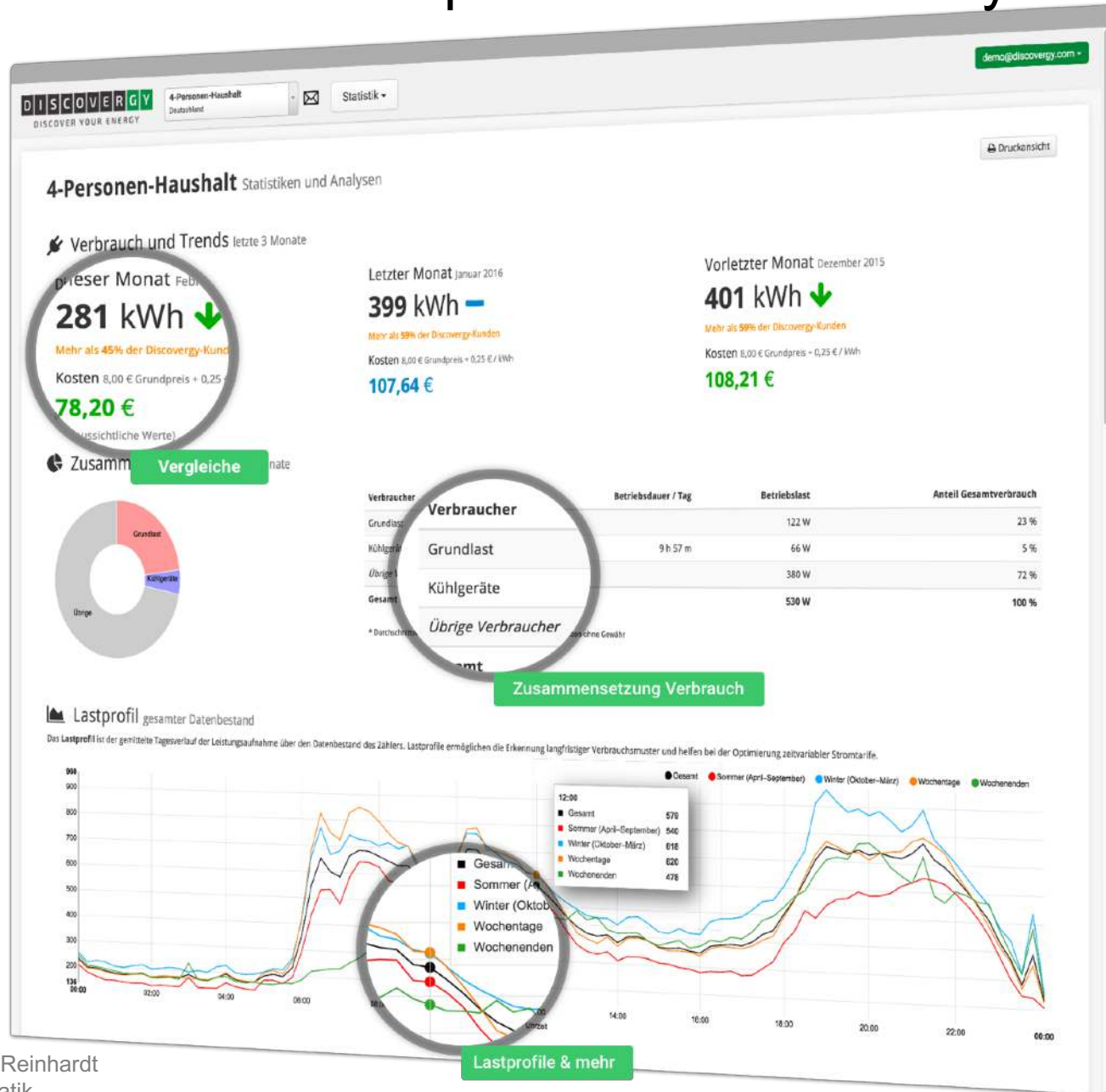
## Stand der Dinge: Auswertung erfasster Daten

(neben der Nutzung zu Abrechnungszwecken)

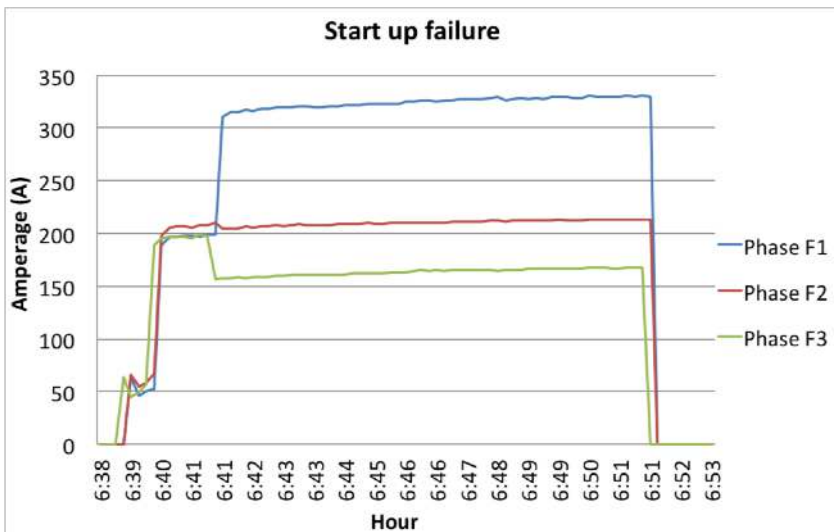
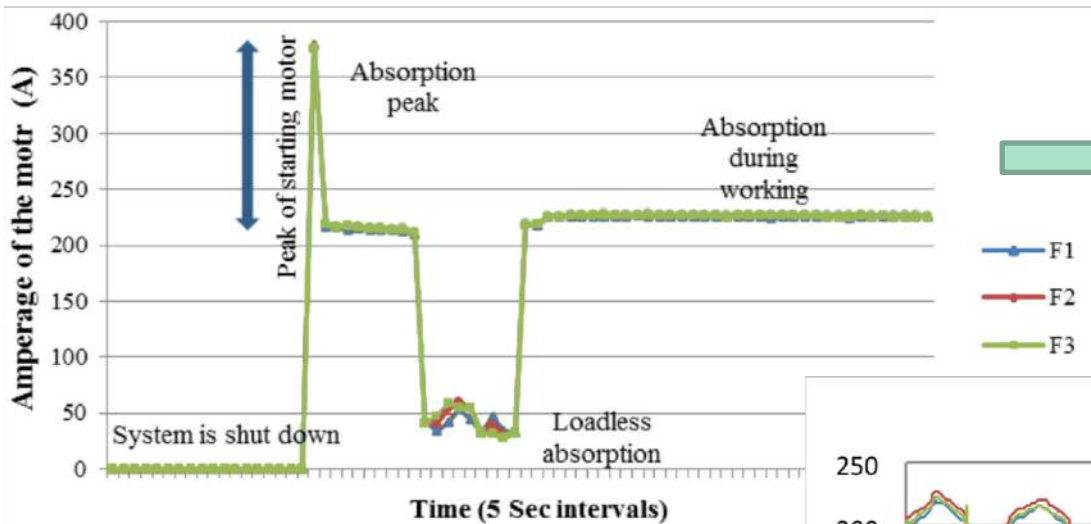
- Im privaten Umfeld
  - Verbrauchsanalyse, z.B. in Form von „Energieberatung“
  - Einsparpotenzial von ca. 5-12% durch Anzeige des Verbrauchs
- Im kommerziellen und/oder industriellen Umfeld
  - Detailliertere Kostenanalyse für den Betrieb von Anlagen
  - Berechnung des Energieverbrauchs für einzelne Prozesse
  - Predictive Maintenance



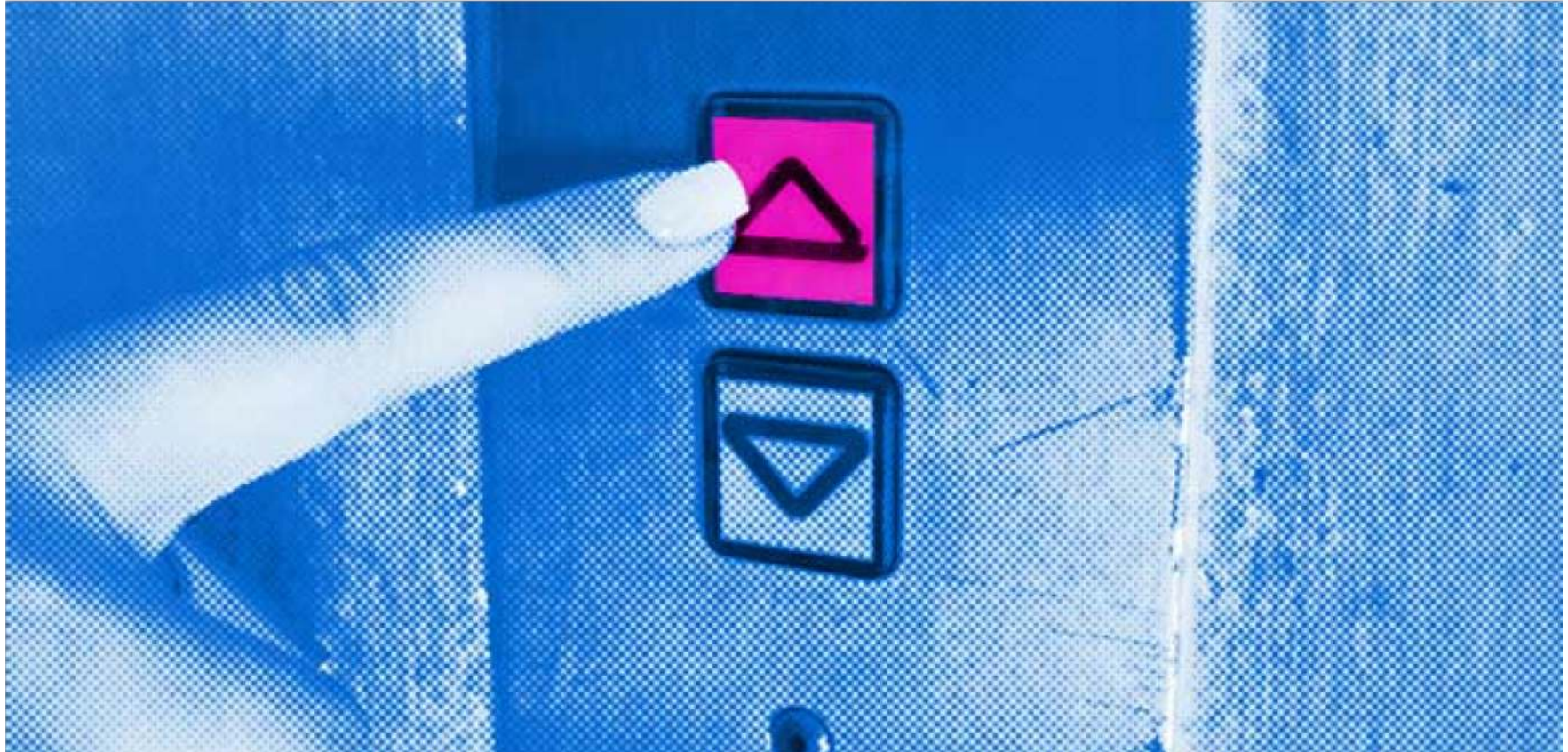




Bildquelle: discovery.com



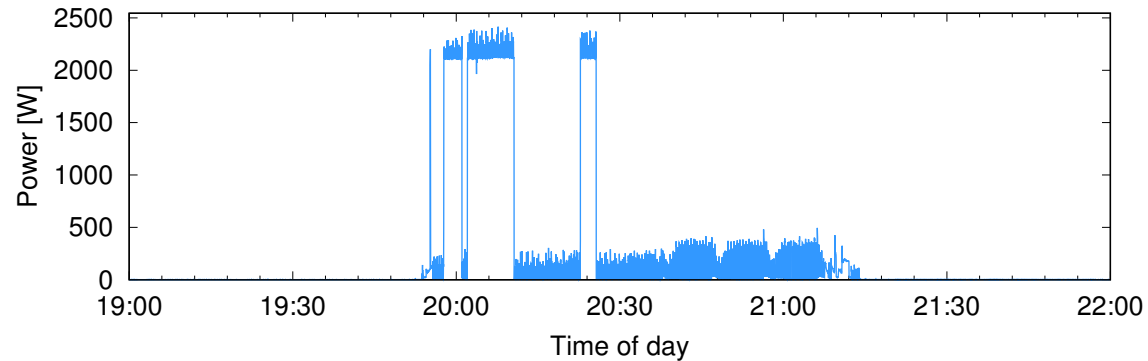
## Auf! Zu neuen Zielen...



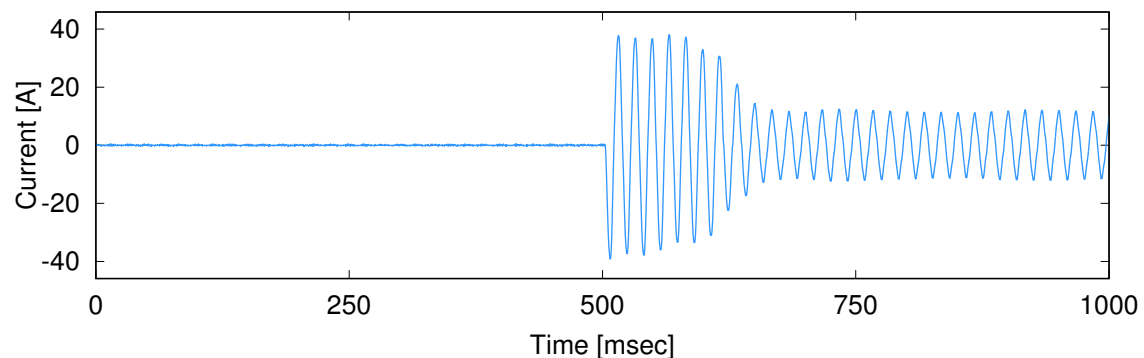
Warum kann der Energieverbrauch erst nach einigen Tagen aufgeschlüsselt werden?  
Warum wurde der defekte Motor erst nach 10 Minuten asymmetrischer Last erkannt?

## Auflösung der zu Grunde liegenden Daten

- *Makroskopische* Datenreihen werden in Intervallen erhoben, die unterhalb der Netzfrequenz liegen (z.B. als Effektivwerte, mit dem damit einhergehenden Verlust an Informationen)



- *Mikroskopische* Datenreihen werden mit Abtastraten erfasst, die deutlich über der Netzfrequenz liegen; dies erlaubt die Analyse der Wellenformen, dafür sind langfristige Trends schwerer zu erkennen



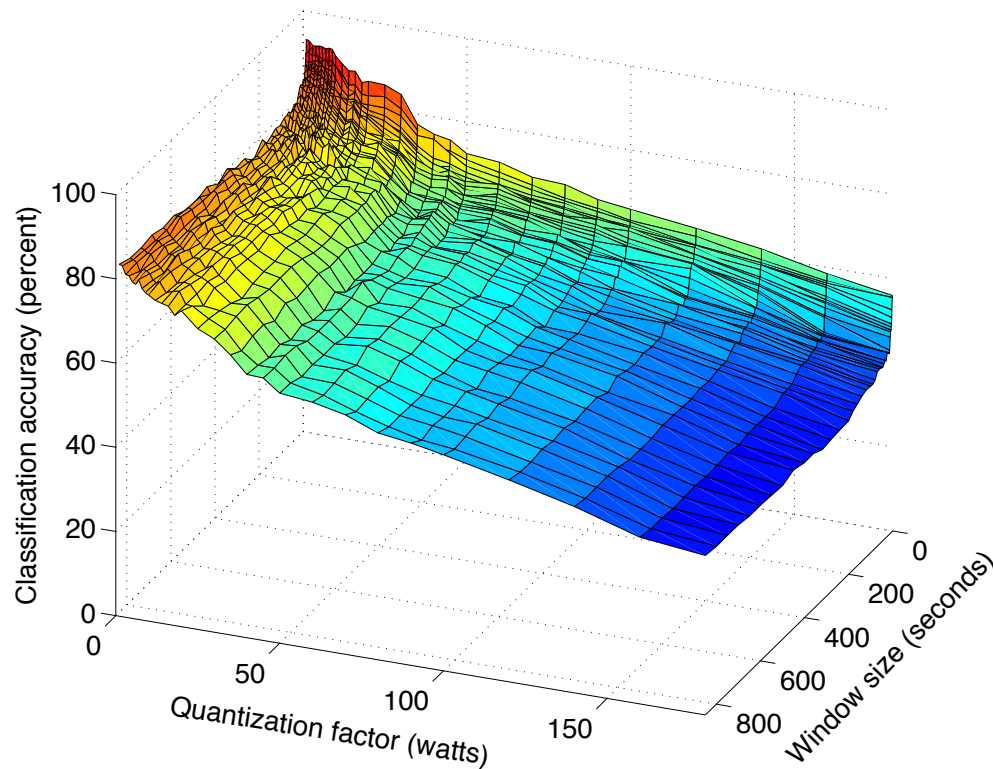
## Datenerfassung führt oft zu großen Datenmengen

- Abhängig von der Abtastfrequenz der Datensätze ergeben sich schnell große Datenvolumina
  - Bei 0,2 Hz (ein Messwert pro fünf Sekunden): 17.280 Werte pro Tag
  - Bei 250 kHz:  $21,6 * 10^9$  (21,6 Milliarden) pro Tag
- Veröffentlichte Datensätze anderer Forschungsgruppen:

Datensatz	Abtastrate	Größe (in Gigabyte)
BLOND	250 kHz (teils 50 kHz)	40566
UK-DALE	16 kHz	4070
BLUED	12 kHz	270
PowerNet	1 Hz	30
PLAID	30 kHz	1,26
REDD	15 kHz	0,83

## Abtastrate vs. Informationsgehalt

- 40 Terabyte? 250kHz Abtastrate? Braucht man das überhaupt?
  - Studie mit makroskopischen Daten (1Hz bis 1 Wert alle 15 Minuten)
  - Aktuelle Forschungsergebnisse deuten darauf hin, dass der Gehalt verwertbarer Informationen bis zu Abtastraten von ca. 1-2kHz weiter steigt



# Beispiele der Verarbeitung von Energiedaten mit Hilfe informationstechnischer Werkzeuge

Bildquelle: [www.projahn.de](http://www.projahn.de)



## Beispiel 1: Erkennung elektrischer Verbraucher

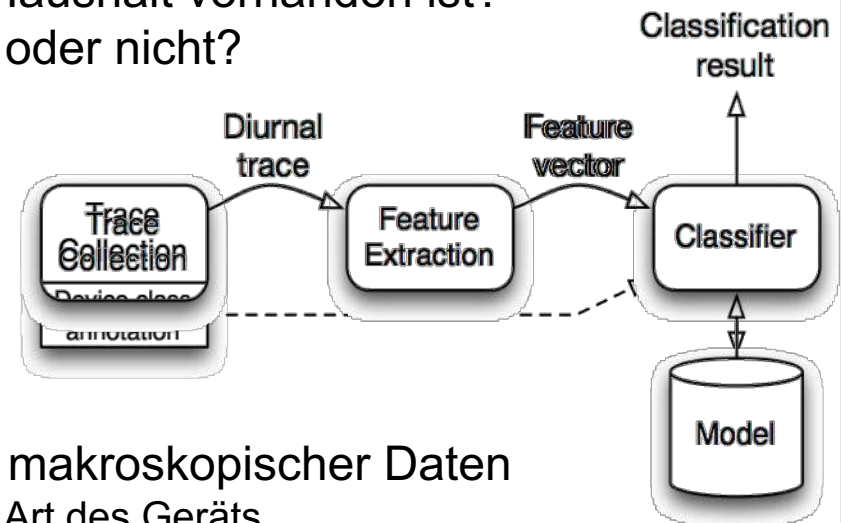
- Kann eine intelligente Umgebung (*smart home*) herausfinden, ob...
  - ...ein bestimmtes Gerät in einem Haushalt vorhanden ist?
  - ...dieses Gerät energieeffizient ist oder nicht?

- Annahmen

- Messgeräte sind vorhanden
- Erkennung darf einen Tag dauern

- Vorgehensweise

- Training auf Basis zuvor erfasster makroskopischer Daten
  - Annotation gesammelter Daten bzgl. Art des Geräts
  - Extraktion von über 500 Merkmalen
  - Anwendung maschineller Lernverfahren und Erstellung eines Modells zur künftigen Erkennung von Lastgängen
- Klassifikation
  - Evaluation des Modells gegen bislang unbekannte Daten
  - Bestimmung der Genauigkeit der Gerätetyp-Erkennung





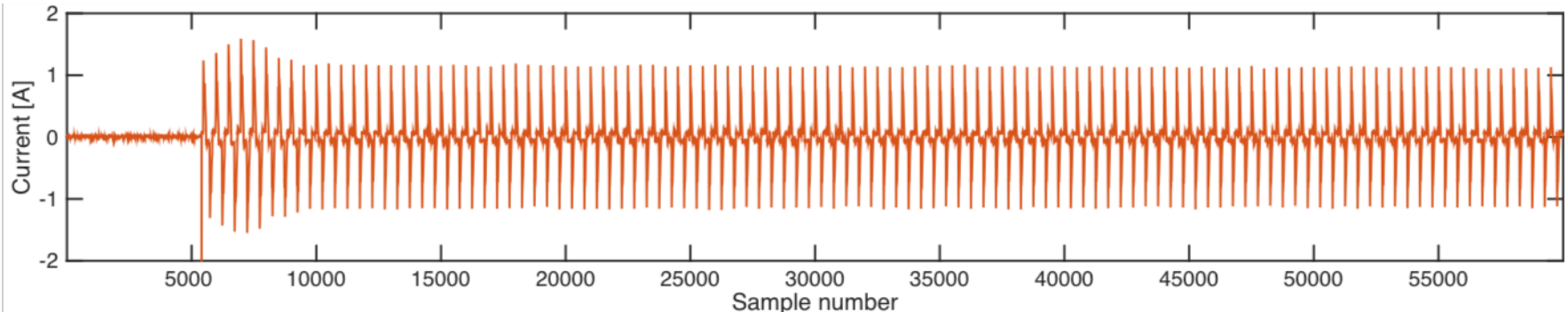
## Auswertung der Daten mittels Konfusionsmatrix

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	1	2	3	4	5	6	7		
A	11	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	
B	0	68	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	
C	0	0	50	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
D	0	0	0	29	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
E	0	0	0	0	20	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
F	0	0	0	0	0	60	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	
G	0	0	0	0	0	0	65	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	1	0	0	0	0	0	0	1	
H	0	0	0	0	0	0	0	40	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
I	0	0	0	0	0	0	0	0	42	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
J	0	0	0	0	0	0	0	0	0	43	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
K	0	0	0	0	0	0	0	0	0	0	31	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
L	0	1	0	0	0	0	0	0	0	0	0	26	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
M	0	0	0	0	0	0	0	0	0	0	0	0	34	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
N	0	0	0	0	0	0	0	0	0	0	0	0	0	4	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
O	0	0	0	0	0	0	0	0	0	0	0	0	0	0	5	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	
P	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	16	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
Q	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	9	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
R	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	12	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
S	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	41	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
T	0	0	0	0	0	0	0	2	1	0	0	0	0	0	0	0	0	0	0	0	38	0	0	0	0	0	0	0	0	0	0	0	0	1	0
U	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	16	0	0	0	0	0	0	0	0	0	0	0	0	
V	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	56	0	0	0	0	0	0	0	0	0	0	0	
W	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	22	0	0	0	0	0	0	0	0	0	0	0	
X	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	3	0	0	0	0	0	0	0	0	0	0	
Y	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	0	13	0	0	0	0	0	0	1	
Z	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	23	0	0	0	0	0	0	0	0	
1	1	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	12	0	108	0	0	0	0	0	0	
2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	126	0	0	0	0	
3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	6	0	0	0	0	
4	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	70	0	0	0	
5	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	16	0	0	0	
6	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	16	0	0
7	0	0	0	0	0	0	2	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	2	1	0	0	0	0	0	35	0

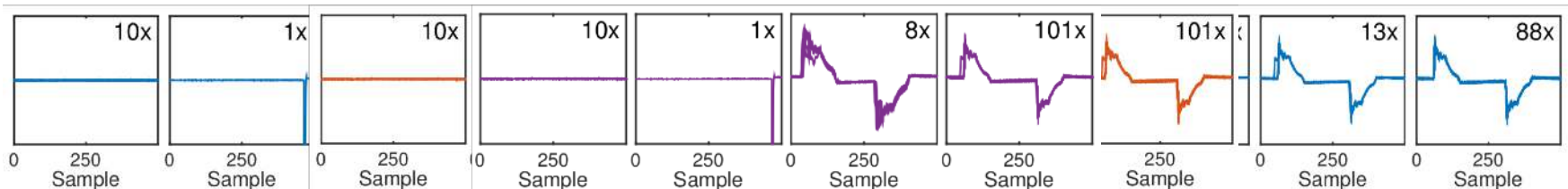
## Evaluation im Feldversuch und Ergebnisse

- Versuchsumgebung und Randbedingungen
  - Training durch Merkmalsextraktion von 1.197 Zeitreihen
  - Einsatz einer 25-fachen Kreuzvalidierung
  - Beim Einsatz eines „Random Committee“-Modells erzielte Werte:
    - Genauigkeit: 95.5% (größte Unsicherheit zwischen LCD- und CRT-Monitor)
    - Zeit zum Erstellen des Modells (exkl. Merkmalsextraktion): 0.42s
  
- Beobachtungen und Schlussfolgerungen
  - Auf Basis sekundlich erfasster Leistungsaufnahmen elektrischer Verbraucher lassen sich Geräte in Haushalten identifizieren
  - Abgleich des Energiebedarfs des erkannten Geräts mit dem Bedarf von Geräten selben Typs → Aussagen bzgl. Energieeffizienz möglich
  - Wissen um das Vorhandensein von Verbrauchern kann aber auch zur „Selbst-Konfiguration“ intelligenter Umgebungen eingesetzt werden
  - Der prototypisch umgesetzte Ansatz beleuchtet bereits den hohen Informationsgehalt von Energiedaten

## Beispiel 2: Anomalien in der Leistungsaufnahme

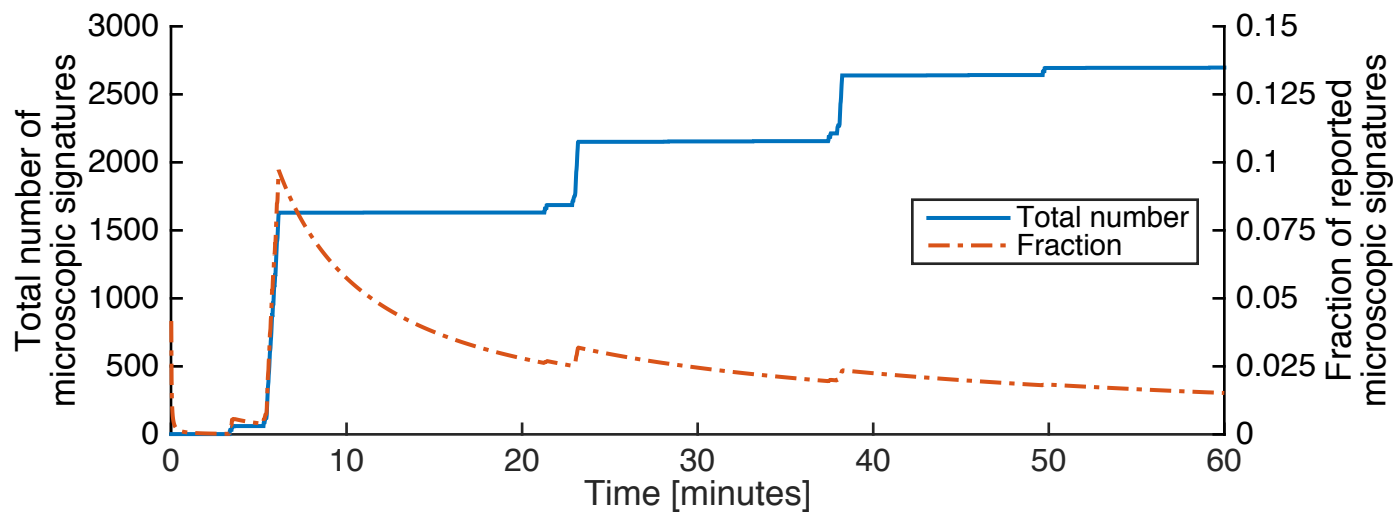
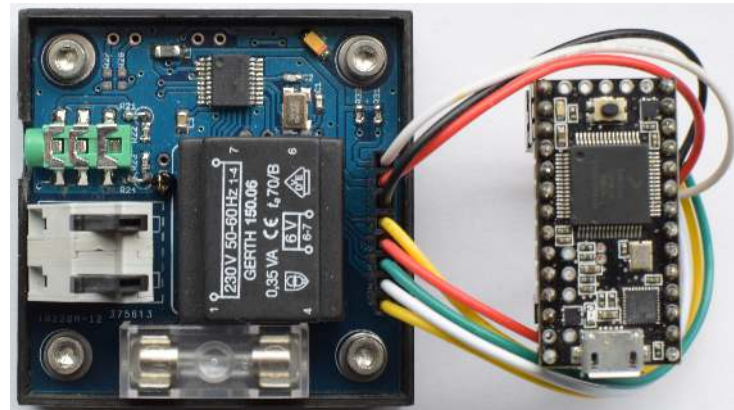


- Betrachtung von 120 Spannungsperioden (d.h. 2 Sekunden) einer Energiesparlampe
  - Quelle: PLAID-Datensatz (Netzfrequenz 60Hz, Abtastrate 30kHz)
- Zerlegung der Wellenform in grundsätzliche Komponenten
  - Bewertung der "Ähnlichkeit" zwischen aufeinanderfolgenden Zyklen



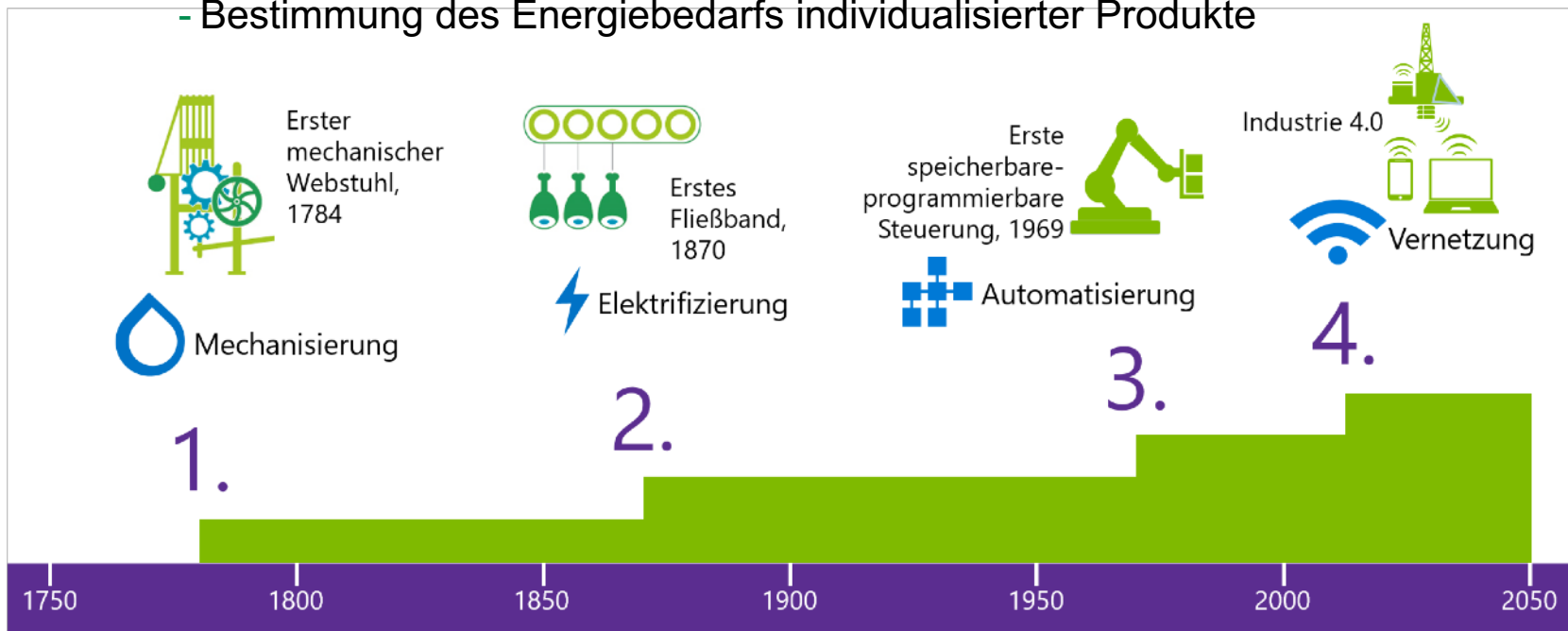
## Studie über einen längeren Zeitraum

- Ausführung des Algorithmus über den Zeitraum einer Stunde
- HP-Lasterdrucker (s/w) an einem prototypischen Mess-System
- Druckvorgang (jeweils fünf Seiten) zu Minuten 5, 23, 38



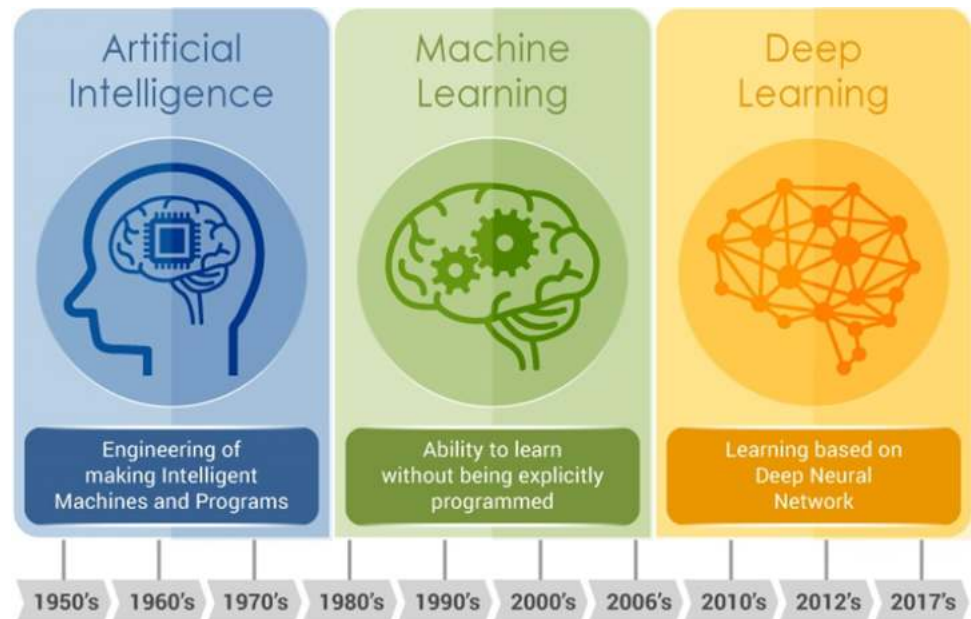
## Einsatzpotenziale im Fokus von Industrie 4.0

- Energiedaten nehmen auch im Kontext von Industrie 4.0 eine wichtige Rolle ein
  - Status-Monitoring, u.a. für „predictive maintenance“
    - Empfehlungen für Austausch oder Reparatur ineffizienter Geräte
  - Korrelation von Energieverbrauch mit anderen Parametern
  - Bestimmung des Energiebedarfs individualisierter Produkte



## Take-home Message

- Tiefgehende Kenntnis von Energieinformationen kann bei der Steigerung der Energieeffizienz maßgeblich helfen
  - Auswahl der richtigen Abtastrate (Datenvolumen vs. Informationsgehalt) und Sensor-Instrumentierung
  - Algorithmen zur Extraktion relevanter Merkmale aus den Daten
  
- Diese können aus dem Bereich der künstlichen Intelligenz stammen oder aus der Signalverarbeitung
  
- ABER: Auch "gute" Algorithmen können bei "schlechten" Daten ihr volles Potenzial nicht ausschöpfen!



# Signalverarbeitung und maschinelles Lernen als Werkzeuge zur Verbesserung der Energieeffizienz

Dr.-Ing. Andreas Reinhardt

