Introduction to Intelligent Process Control-Seminar:
“what to expect from AI technologies”

Holger Schlarb from DESY Hamburg, 11\textsuperscript{th} of Aug. 2020

Outline:
- Introduction & goal of IPC-seminar
- Recap on our accelerators and perspectives
- AI & ML technologies
My background

issues …

- Studied physics Ffm & HH with some detour on math. (num. PDE/Distr. Theory)
- MSc & PhD in Accelerator Physics in Hamburg (Wakefield comp. / Collimation for TTF)
- 2003 Sabbatical SLAC (SPPS commission)
- 2005 XFEL-WP18 Leader “Special Diag.” → 2010 Ch. Gerth
- 2010 XFEL-WP02 Leader “Low Level RF controls” → 2012 J. Branlard
- 2011 Head of “Accelerator Beam Controls” Group (MSK)
- 2017 Involved in DESY’s scientific computing strategy discussion

Domain Expert in acc. Physics

Disclaimer: I am not a computer scientist & not at all in AI !!!!

Studied quite a bit (mainly regarding methodologies / technologies)
But: little practical experience
Still in learning phase….
Scientific computing @ M-division
What already exists … & we are pretty strong at

• Numerical codes for beam dynamic simulations…(ASTRA, CSRTrack, …)
• Theory & Simulation of FELs (FAST, Genesis, ALICE)
• Electromagnetic fields simulations (e.g. ECHO) + coop. with TEMF (Darmstadt)
• Fast tracking & opti. codes for operation (RFTweak, OCELOT, …)
• Plasma Wakefield Acceleration …. (HPC required)

➡ Dedicated Sci. Comp. groups in M-Division (well covered)
★ seminars: e.g. Beam Dynamics XFEL/PETRA/PWA (Igor/Mikheil/Ilya/Max)
Focus of the IPC-seminar:

Why a new seminar?

Intelligent Process Controls…

► Scientific Computing related to

- Data mining and data analysis
- Advanced and data driven controls
- Fault diagnosis and supervisory control
- Automated root cause analysis / causal inference
- Sophisticated automation routines
- Information extraction technologies
- Advanced statistical methods and optimization algorithm
- Explore modern AI methods and technologies

applied to accelerator design, construction, commission and operation

Has both:

Methodological & Technological aspects!

Ties on different scientific computing field as we are used to.
Interdisciplinary work essential …

Both have to learn….

- Different language, notation and annotation + pre-existing knowhow
- Different approaches, viewpoints and methodologies
- D-Expert: limited method overview / adequate implementation / structural analysis / identify use cases
- SC-Expert: formulation of goals, system design & interconnect, may own agenda (method)

→ Buildup knowhow on the respective other domain (specially language)
→ Buildup teams from both domains
Goal of the IPC-seminar:

- **IPC-seminar should serve**
  - Bringing the acc. domain expert and computer science experts together
  - Communication platform for all types of SC problems \( \rightarrow \) explore synergies / joint tool development
  - Add tutorial style talks for education and knowledge dissemination
  - Platform to get know to one another (who is working on what with which methods & expertise)
  - Opportunity to present your results or describe your problems
  - Open minded, low barrier … questions and suggests are welcome … new field for many of us

- **SC@M: concerns M-division as whole / can provide from:**
  - Across all accelerator & laser system & test facilities
  - Across M-division groups
  - Currently virtual group in M-division with members in different groups
Intelligent process control of distributed systems

... requires wider expertise ...

Intelligent Process Controls

- Control System Architecture
- Software Quality
- Metadata
- Data Analytics
- AI & Machine Learning
- Controls Theory
- Fault Diagnosis
- Accelerator Expertise
- HPC
- Hardware Technology
- Streaming Technologies
- Data Base Technology
- Technology & Software Engineering
- Machine-human interface
- Mathematical Methodology Modelling

Personnel ~ 15-20 head counts
LKI/LKII & 3rd party funds
Data and computing science – strategy update
First steps towards a DESY wide scientific computing strategy

Third party funded activities and collaborations

- LEAPS
- CDCS/UHH
- Helmholtz Artificial Intelligence
- Helmholtz Imaging Platform
- DASHH/HIBRIDS graduate schools
- Helmholtz Cross Cutting Activity / HPC
- Universities
- Campus partners

DESY applied scientific computing services

- Big data HPC
- Visualisation
- Machine learning
- Data engineering

Core services in central IT (HPC hardware, storage, network)

DESY MT/DMA

- FH LK I, LK II
- FS LK I, LK II
- M LK I, LK II
- AP LK I

SciComp FH
SciComp FS
SciComp M
SciComp AP

- W3: Big data HPC
- W3: Inv. problems, Imaging (HIP)
- W3: Machine learning
- W3: Data engineering

Domain expertise required to link to science applications
Research in the most important computational fields (academic backing)
Cross-divisional applied scientific computing expertise (service)
Shared base infrastructure

DMA = Data Management Analysis

DESY. Edgar Weckert | Directors Talk | 9 April 2020
Short recap:

DESY’s Accelerators & test facilities
DESY accelerators & test benches: birds view …

Synchrotron radiation source (highest brilliance)

VUV & soft-x-ray free-electron laser

X-Ray Free-Electron Laser
atomic structure & fs dynamics of complex matter
Different accelerators … different challenges

Accelerators at DESY …

FELs… (reproducibility / optimization / flexibility / …)

Storage ring
(lattice / beamlines / optic & orbit / injection / availability/…)

Equilibrium but non-linear

No equilibrium

Highly non-linear

Plasma accelerators (lasers control/ HPC /in-situ FB)

Challenges and degree of matureness very different, but
→ Increased complexity of controls
→ Higher demands on accelerator operations
→ Push on the limits regarding performance & flexibility & availability
Beam dynamics of an FEL drive beam

issues …

• **First of all:** it is a linear accelerator! → no equilibrium state
  → every shot is different
  
  High control effort!!!

• **Photo injectors:**
  • Complex beam transport (up to second order)
  • Bunch compression
  • Collective effects
    • Space charge
    • Coherent Synchrotron Radiation
    • Wake fields
  • Micro bunching instability …
  • … many more

![Beam dynamics diagram](image-url)
Beam dynamics of an FEL drive beam

Self-Amplified Spontaneous Emission (SASE)
(Kondratenko, Saldin 1980)
(Bonifacio, Pellegrini 1984)

\[ \lambda_{ph} = \frac{\lambda_u}{2\gamma^2} \left( 1 + \frac{K^2}{2} \right) \]

Resonance wavelength:

Undulator parameter \( \approx 1 \)

- Beam parameters challenging to achieve
- Control over 6D phase space mandatory
- kA peak currents causing strong collective effects

Stochastic process

L \sim 200m
Wide dynamics range & different operation modes

Time structure & multiple beamline (SA1/SA2/SA3)

- **Pulsed operation (RF ~ 1.4ms, 10Hz)**
  - \(\Rightarrow\) 27000 / sec
  - \(\Rightarrow\) e- bunches 220ns spaced
  - 100ms separation

- **Bunch distribution & pattern**
  - Fast Kicker System
  - Precise (Slow) Flat-Top Kicker
  - 5/10Hz splitting

- **Change of beam properties for different beam lines**
  - Transients processes
  - Subsystems in different states
  - Large reconfigurations

Challenge:
3 machines in one….. Coupling….
Large and complex Infrastructure & Subsystem

Infrastructure systems (Cryo/Cooling/Power) & large and diverse components ~ 10000 typically

Large number of sensors & components

Large supply infrastructure: Cooling, Systems/Power Distribution/Climateisation/Cryo/Infrastructure...

Here at the example of the Cryoopen-System:

- 675 vertical sensors
- 1967 temperature sensors
- 859 pressure sensors
- 21 individual sensors
- 1200 level transmitters
- 435 regulation loops
- x 23300 meters
- x 233000 properties

Similar complexity for water-cooling & power distribution...

Large number of sensors & components

For beam diagnostics ... mostly bunch-to-bunch resolution required

Installed/Diagnostic items

Number

Beam Position Monitors
Charge Monitors
Imaging Stations
Dark current monitors
Pipe Scanners
Loss Detectors
Discreteness Systems
Transverse Deflecting Structures
Bunch Compression Monitors
Beam Arrival Time Monitors
Electro-Optical Systems
THz spectrometer

Large number of sensors & components

Beam line devices...

Magnets
- 161 Dipole, 495 Guide, 59 Multipole, 493 Correctors, 103 Quad-Waves, 2 Solenoids

Modules
- 101 x 1.0 GHz, 1 x 3.0 GHz, 27 RF Stations
- 501 SKP cavities & RF Coupler & ROIs...

Foil Devices
- 34 kicker magnets, 3 transverse deflecting structure

Undulator
- 1 Laser-Heater Undulator, 91 SASE Undulator Segments

Vacuum
- Total 4480 m of cold & warm vacuum in various sections beam (1), coupler (99), laser (21), dump (5); ion (21) adds to 763 spallation neutron pumps / 40 ion pumps / 54 valves / 8 feedthroughs / vacuum pumps

High power beam & beam density

Machine protect to prevent acc. & photon sub-system damages

- 185 collectors boards
- 4500 different signals (magnets /AC/DC loads / vacuum / diagnostics / optical & acoustics)
- up to 4000 / 4000 beam power
- Damage when ~ few W or possible

Infrastructure & Diagnostics

Magnets & RF

Machine Protection
Changes due to digitalization...

Driven by performance requirements and needs for higher flexibility...

Very complex sensor technologies …

& analog / digital processing units …

...where next evolution is already on the way (Network on Chip = NoC)

Separation from human controls …

- Configuration  data quality
- Demand on supervisory systems
- New HW technologies
What changed over the past decade...

Digitalization changes …

- Classical 3 tier structure + in cooperated Data Acquisition System at EuXFEL

  - > 10 million control parameters
  - > 700,000 local archives
  - > 20,000 high data rate channels
  - > 30 TB/day DAQ (compressed)

  → Configuration management already challenging

  → Large amount of data available
    - data mining
    - information extraction
    - machine / system optimization
    - failure detection
    - predictive maintenance
    - …

  → < 1% sent from front-ends

  "We would like to better make use of this data"

  "tese: T. Wilksen"

- Requires new techniques & technologies → AI/ML & automation
- Integration into controls architecture → needs carefully to be reviewed

Standardize hardware & software

MicroTCA .4

Desy

Intelligent Process Control Kickoff Meeting | Holger Schlarb | Hamburg, IPC-Seminar 11.08.2020
What makes these machine complex …

Challenge caused by combination of performance / flexibility & availability…

Particle accelerators are …
- Largely distributed …
- Various type of systems …
- Highly non-linear processes …
- Strongly coupled (sub-)systems …
- Timing varying disturbances …
- High dimensionality ~ few 1000
- Use of cutting-edge technologies… (sensors / front-end electronics / real-time processing)
- Limited access to key observables

Facing:
Increasing demand on **Performance, Flexibility & Availability**

- Bunch pattern (flexible, on demand)
- Switching bunch pattern modes (10Hz)
- Multi-beamline operation
- Fast tuneability within trains
- Special modes
- …

- Ideally 99%
- Small setup times (reproducibility)
- Small tuning times (operability)
- Reduced scheduled down
- …

~ μm x-y orbits & <dispersion
~ mdeg / 0.01% RF phase/ampl.
~ fs arrival time @ km scale
~ mK temp. stability in parts
~ 30 ubar 2K pressure
~ beam parameters ($\lambda, \sigma_t, E_{ph}$, chirp..)

...
Future operation & future machines .... perspectives

Questions

- can we tailored beams with high precision for user experiments
- how far can we push accelerator stability and reproducibility
- how can we predict control over sufficiently long time horizons
- can we develop autonomous accelerator operation
- can we establish new user-modes to scan the right physics
- can we build accelerators that predicts failures and repair itself
- can we enable new accelerator designs not accessible today

by applying Advanced Controllers, Data Analytics and AI technologies for intelligence process control of accelerators
Artificial Intelligence...
What is artificial intelligence …

Some definition …

Artificial intelligence is the attempt to simulate rational or cognitive human intelligence on (technical) machines in order to use it in a profitable and beneficial way for humans

Classification:

1. Appropriate intelligence - adequate reaction to environmental stimuli (I1)
2. Learning intelligence - independent acquisition of new knowledge (I2)
3. Creative intelligence - generation of knowledge outside of formal induction and deduction (I3)
4. Conscious intelligence - Conscious understanding of knowledge (I4)
5. Self-confident intelligence - conscious understanding of the “ICH” concept (I5)

Example I1:

Example I2:

Technical AI systems that are able to apply complex knowledge and at least partially autonomously generate new knowledge

- speech processing systems,
- image recognition systems,
- expert systems,
- rule-based systems,
- automatic evidence systems,
- …

Remark: AI-definition in Europe (how its done) differs from Anglo-Saxon (what can it do)
Methodology to gain knowledge

Some definition …

**Deduction:** ("Inferring Forward", "Applying Forward" a rule base)
- Premise (axiom) + inference rules (known knowledge) → rules (new knowledge)
- Cause (observed facts) + rules of conclusion (known knowledge) → effect (new facts)

E.g. Statement logic, predicate logic PL1/2, fundamental nature / engineering laws, math. construction..

**Abduction:** ("inferring backwards", "applying backwards" a rule base)
- Effect (observed facts) + rules of conclusion (known knowledge) → cause (new facts)

**Induction:** (»learning from data«, creating a non-generally applicable rule base)
- Cause (observed facts) + effect (observed facts) → rule (new statistical knowledge)

**Remark:** in math. full induction is correct since complete!

Always incomplete

Safety systems: NO GO!

- Only of statistical nature
- Cannot extrapolate! ×
- Generalization … good interpolation
AI landscape

A possible way of categorization ...

<table>
<thead>
<tr>
<th>Intelligent Process Control Kickoff Meeting</th>
<th>Holger Schlarb</th>
<th>Hamburg, IPC-Seminar 11.08.2020</th>
</tr>
</thead>
</table>

Supervised: Y exist
Un-supervised: no Y
Reinforced: receives reward

Big hype:
Knowledge Discovery in Data Base (KDD)

Electronic analog circuits to mimic neuro-biological architectures
What can you do with data using ML

issues …

1. Columns: \( G = f(A, C, E) \) ask for functional relation (regression) variable to each other? → directed relation

2. Rows: how many clusters exist? → structural analysis

3. Individual rows: affiliation to a cluster? → structural analysis

4. Individual columns to one-another: correlation analysis → undirected relation

5. Succeeding rows: prediction for following rows → trend analysis
ML: first think you learn …linear regression…  
which is kind of standard for natural scientist & engineers…

Best estimate for $y = f(x)$:

$w_0 + x_{1,p}w_1 + x_{2,p}w_2 + \ldots + x_{N,p}w_N \approx y_p, \quad p = 1, \ldots, P.$

Columns: input features $x_n$

Bias

Feature touching weights $w_n$

Least Square cost function $g(w)$:

Minimize $\frac{1}{p} \sum_{p=1}^{P} (\hat{x}_p^T w - y_p)^2$  

Metric required (here just $l_2$: $p \parallel \cdot \parallel_{l^2}^2$)  

$\triangledown g(w) = 0 \Rightarrow \left( \sum_{p=1}^{P} \hat{x}_p^T \hat{x}_p \right) w = \sum_{p=1}^{P} \hat{x}_p y_p$  

Can be solved in one Newton step

$\Rightarrow$ Local optimization problem (convex cost function)

$H_{ij}(g(w)) = \partial_i \partial_j g(w)$

Remarks:

- Concept easily extends to non-linear regression (add new features $x_i^n x_j^m$)  
  $\Rightarrow$ if feasible preferred way, coefficients have defined meanings

- $l_2$ norm overfits outliers $\Rightarrow$ $l_1$ “last absolute deviation” $| \cdot |$

- For larger $N$ … gradient descents method used to compute $w$
ML: sort detour to “Optimization Problems”
most cases finally end up as…

General formulation:

**nonlinear program**

\[
\begin{align*}
\inf_z & \quad f(z) \\
\text{subj. to} & \quad g_i(z) \leq 0 \quad \text{for } i = 1, \ldots, m \\
& \quad h_j(z) = 0 \quad \text{for } j = 1, \ldots, p \\
& \quad z \in Z,
\end{align*}
\]

\[f : \mathbb{R}^s \rightarrow \mathbb{R}, \quad g_i : \mathbb{R}^s \rightarrow \mathbb{R}, \quad h_i : \mathbb{R}^s \rightarrow \mathbb{R} \]

**Convex:**

\[
\lambda z_1 + (1 - \lambda) z_2 \in S \\
\frac{\partial}{\partial z} f(\lambda z_1 + (1 - \lambda) z_2) \leq \lambda \frac{\partial}{\partial z} f(z_1) + (1 - \lambda) \frac{\partial}{\partial z} f(z_2)
\]

Global optimum \( f(z^*) = f^* \)

**linear program**

\[
\begin{align*}
\inf_z & \quad d^T z \\
\text{subj. to} & \quad G z \leq w \\
& \quad A z = b,
\end{align*}
\]

\( G \in \mathbb{R}^{m \times s}, \quad w \in \mathbb{R}^m \\
A \in \mathbb{R}^{p \times s}, \quad b \in \mathbb{R}^p \\
\)

**quadratic program**

\[
\begin{align*}
\min_z & \quad \frac{1}{2} z^T H z + q^T z + r \\
\text{subj. to} & \quad G z \leq w \\
& \quad A z = b.
\end{align*}
\]

\( z \in \mathbb{R}^s, \quad H = H^T \geq 0 \in \mathbb{R}^{s \times s}, \quad q \in \mathbb{R} \)

Unbound: \( z^* = -H^{-1} q \)

Unbound: \( z^* = -H^{-1} q \)

H^{-1} computational expensive \( \rightarrow \) gradient method

( obviously high order methods are more efficient)

Cost function \( f, g, h \) might be explicit, implicit or as oracles

Constrains

Goal to find \( z^* \) “the optimizer”

- Global optimum \( f(z^*) = f^* \)

- \( \infty \)

- \( \lambda \)

- \( z \)

- \( H \)

- \( q \)

- \( b \)

- \( G \)

- \( A \)

- \( w \)

- Convex:

- Global optimum

- Instable

- Converges
ML: Nonlinear Feature Engineering…

First steps towards universal approximators…

Inputs (features) \[ \text{But use now non-linear fct.} \]

\[
\text{model} (\mathbf{x}, \Theta) = w_0 + f_1 (\mathbf{x}) w_1 + f_2 (\mathbf{x}) w_2 + \cdots + f_B (\mathbf{x}) w_B
\]

Represent all possible parameters (weights), also inside fct.

- Wavy appears \( \Rightarrow \) use trigonometric functions

\[
\begin{align*}
    f_1 (\mathbf{x}) &= \sin (v_{1,0} + v_{1,1}x_1 + v_{1,2}x_2) \\
    f_2 (\mathbf{x}) &= \sin (v_{2,0} + v_{2,1}x_1 + v_{2,2}x_2)
\end{align*}
\]

- Elliptical boundaries

\[
\text{model} (\mathbf{x}, \Theta) = w_0 + x_1 w_1 + x_2 w_2 + x_1 x_2 w_3 + x_1^2 w_4 + x_2^2 w_5.
\]

- Neural network family \( \quad f_b (\mathbf{x}) = \tanh (w_{b,0} + w_{b,1}x) \)

Generalization:
Degree of freedom &
Structure correctly chose
ML: clustering methods

... for data structure analysis, outlier identification, ...

Basic sequence of a cluster analysis

Classification scheme

> Likely dedicated talk/tutorial on clustering methods

DESY | Intelligent Process Control Kickoff Meeting | Holger Schlarb | Hamburg, IPC-Seminar 11.08.2020
Artificial neuronal networks as universal approximators

Basic intro...

- Single neuron: algebraic represented by
  \[ f^{(1)}(x) = a \left( w_0^{(1)} + \sum_{n=1}^{N} w_n^{(1)} x_n \right) \]

- 1 hidden layers: algebraic represented by
  \[ \text{model}(x, \Theta) = w_0 + f_1^{(1)}(x) w_1 + \cdots + f_{U_1}^{(1)}(x) w_{U_1} \]
  \[ f_i^{(1)}(x) = a \left( w_{0,i}^{(1)} + \sum_{n=1}^{N} w_{n,i}^{(1)} x_n \right) \]
Artificial neuronal networks as universal approximators

Basic intro...

- **Multi-layer neural networks:**
  \[ \text{model } (x, \Theta) = w_{L+1}^T \hat{a} \left( w_L^T \hat{a} \left( \cdots \hat{a} \left( w_1^T \hat{x} \right) \right) \right) \]
  - concatenated chain of multi-dim. function
  - allows to approx. arbitrary functions

- **Specialties of ANNs**
  - Structure: parallel distributed processing
  - Natural propensity for storing experiential knowledge

- **Break through:**
  - Efficient training methods (initialization, regulation techniques, batch-training, dropout, efficient optimizers, avail. libraries....)
  - **GPUs:** massive improvement on parallel processing
  - **Remark:** large ANN still takes long time to train

- **ANN strength:** Enhancement due manipulation of sub-symbolic elements
- **ANN criticism:** Missing interpretability of sub-symbolic elements

Example full-connected neural network for two-class dataset
Some special types of ANN:

Basic intro...

- **Convolutional NN:**
  - Mainly of image processing
  - Feature maps using convolution

  ![Convolutional NN diagram]

  \[ S[i, j] = (K * I)[j, j] = \sum_{m} \sum_{n} I[i + m, j + n] \cdot K[m, n] \]

  Typically 3x3 / …/11x11

- **Specialty:**
  - learns kernel function for feature maps
  - Sparse interactions
  - Parameter sharing
  - Equivalent representations

  ![Typical architectures diagram]

  ![Feature maps and kernels diagram]

  Ref: Hinten2017 DOI:10.1145/3065386

@Book by Frochte
Some special types of ANN:

Basic intro…

- **Autoencoder-NN:**
  - For feature reduction / extraction.
  - Target value is input value $y = x$.
  - Ask for a function of type
    $$x = f_{\text{dec}}(f_{\text{enc}}(x))$$

- **Specialty:**
  - Use for data reduction
  - Noise removal
  - Non-linear dimension reduction
  - Decoupling of learning cycles (weak hardware)
Some special types of ANN:

Basic intro…

- **TM-Polynomial-NN:**
  - Activation functions are polynomials
  - E.g. solving non-linear ordinary differential equations
    
    \[
    \frac{d}{dt}X = F(t, X) = \sum_{k=0}^{\infty} P^{1k}(t)X[k]
    \]
    
    F as Taylor expansion
  - Evolution of system from \( t_0 \to t \) expressed by Matrix Lie maps
    
    \[
    X_{i+1} = M \circ X_i = W_0 + W_1 \begin{pmatrix} x_i \\ y_i \end{pmatrix} + W_2 \begin{pmatrix} x_i^2 \\ y_i^2 \end{pmatrix} + W_3 \begin{pmatrix} x_i^3 \\ y_i^3 \end{pmatrix}
    \]
  1) ODE → calc. W → efficient numerical solver for ODL
  2) Trajectory → train W → extract ODL

- **Specialty:**
  - Shallow but very deep nets
  - Clear interpretation of weights
  - Trick: Regularization includes symplecticity of \( W_i \)

Talk by Andrei Ivanov
To wrap up

- **Digitalization** and **complexity increase** of accelerator $\rightarrow$ new tools for controls required

- **Supervisory control** of processes mostly requires to build **models** of your systems:
  - **AI & ML** (I2): allows to construct models / rules from data, but they are of **statistical nature**
  - **Strength of ANN**: learning and manipulation of sub-symbolic elements derived from features
  - **Data Mining**: make use of our data base for better understanding of our systems on all levels
  - **Predictive controls**: we like to achieve $\rightarrow$ prediction of accelerator behavior for finite time horizons
  - **Goal of seminar**: bring to together the various domain experts across M-division

![Diagram showing White box, Black box, Grey box categories with their respective methodologies.](image-url)
Some additional information

- **Machine Learning at DESY**: desy-ml@desy.de
  Indico page: https://indico.desy.de/indico/category/641/

- **OWLE**: Virtual seminar series on computational accelerator physics, machine learning and experimental demonstration of AI-ML; [https://sites.google.com/view/owle/](https://sites.google.com/view/owle/)

- **IPC – mailing list**: ipc-info@desy.de
- **IPC – confluence**: [https://confluence.desy.de/display/IPC](https://confluence.desy.de/display/IPC)
- **IPC – seminar**: [https://indico.desy.de/indico/category/740/](https://indico.desy.de/indico/category/740/)

---

**General Contact:**
- Annika Eichler
- Raimund Kammering
- Ilya Agapov
- Holger Schlarb
Thanks for attention
Literature:

- *Maschinelles Lernen: Grundlagen und Algorithmen in Python*

- *Predictive Control for Linear and Hybrid Systems*
  - ISBN 978-1-107-01688-0

- *Quantum Computing verstehen: Grundlagen - Anwendungen - Perspektiven*

- *Künstliche Intelligenz für Dummies*
  - ISBN: 978-3-527-71494-0

- *Model Predictive Control*

- *Hands-on Machine Learning with Scikit-Learn & TensorFlow*
  - ISBN 978-1-491962299

- *Machine Learning kurz & gut*
  - ISBN 978-3-96009-052-6

- *Causal Inference in Statistics: A Primer*

- *Reinforcement Learning Algorithms with Python*

And many more….