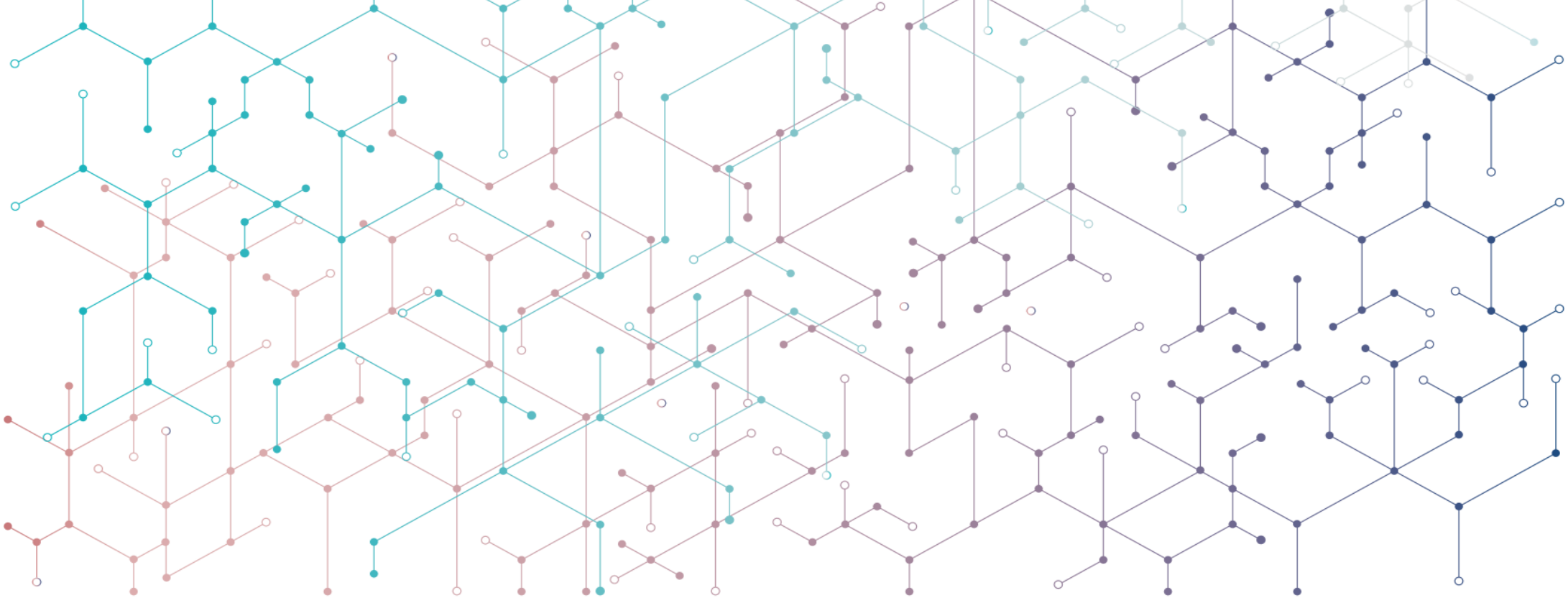


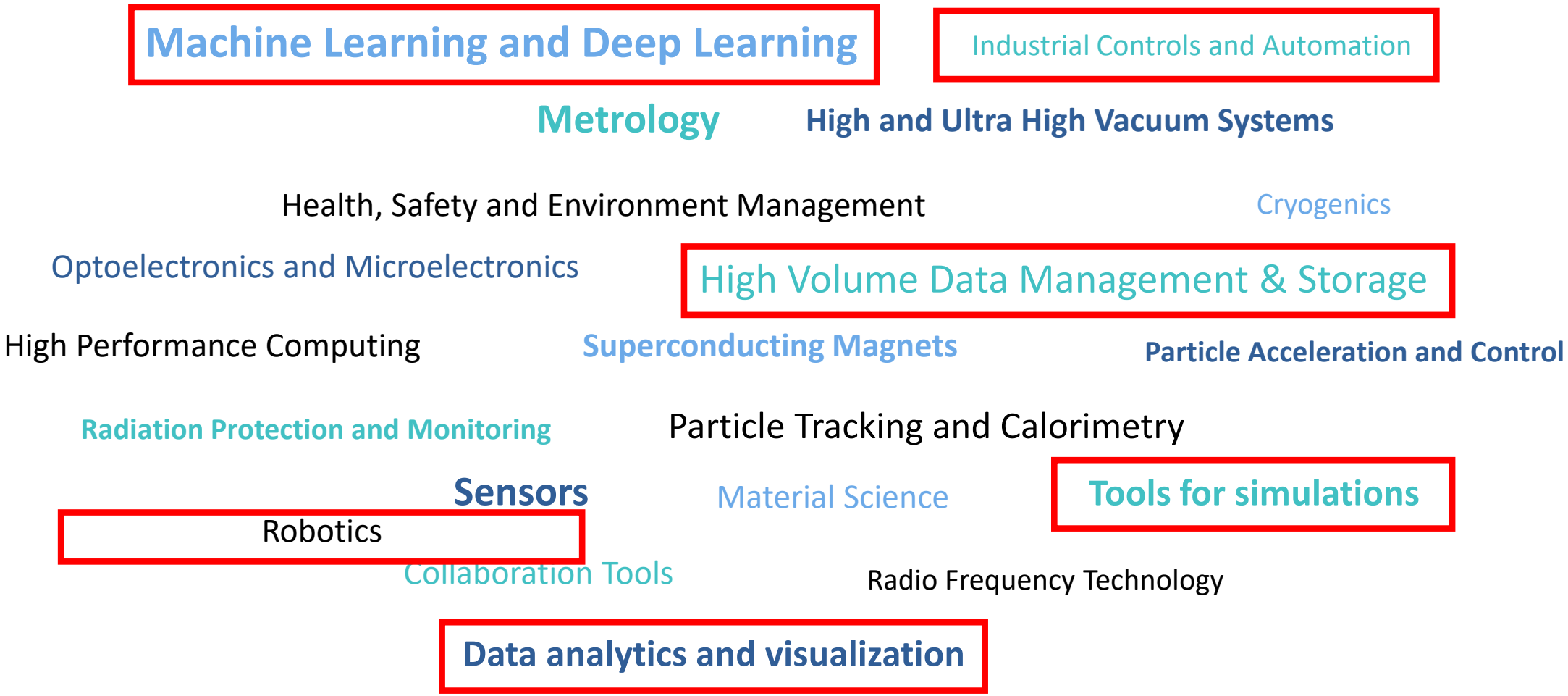
Working with CERN Knowledge Transfer

Sample of examples cases and projects

Nick Ziogas
nick.ziogas@cern.ch
Open Lab Technical Workshop 22 March 2022

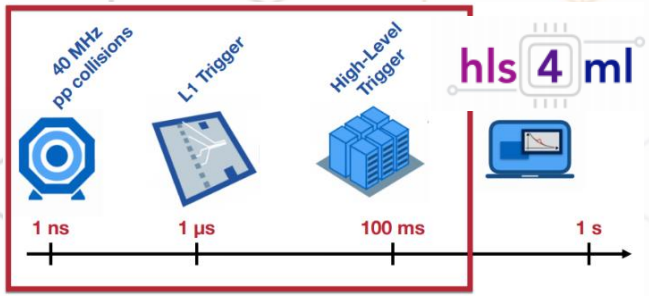
Our Mission

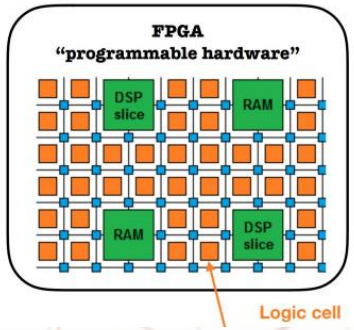
- **Maximise** the technological and knowledge return to society, in particular through Member States industry
- **Promote** CERN as a centre of excellence for technology and innovation
- **Demonstrate** the importance and impact of fundamental research investments


Machine Learning and Deep Learning
Industrial Controls and Automation
Metrology
High and Ultra High Vacuum Systems
Health, Safety and Environment Management
Cryogenics
Optoelectronics and Microelectronics
High Volume Data Management & Storage
High Performance Computing
Superconducting Magnets
Particle Acceleration and Control
Radiation Protection and Monitoring
Particle Tracking and Calorimetry
Sensors
Material Science
Tools for simulations
Robotics
Collaboration Tools
Radio Frequency Technology
Data analytics and visualization

Key technology: Ultra-fast on-edge neural network inference

Know-how in neural network pruning and neuromorphic chips (e.g. FPGA) for ultra-low latency, on-edge inference





What

CERN needs ultra fast deep learning inference **(execution in ~1 microsec)** for fast classification of particle collision data, requiring compact code for edge-computing on programmable chips (FPGAs with logic cells)

For this, CERN contributes to an open-source **package (hls4ml) to automatically translate pre-trained neural networks** (as specified by NN architecture, weights, biases) into high-level synthesis code for FPGA architecture, drastically **accelerating prototyping, reducing time to results**

Apps

- A **car manufacturer** teamed up with CERN to develop ML for fast classification of image (computer vision) and lidar sensor data from self-driving cars using FPGAs
- A **developer of wireless and multimedia tech** jointly with CERN developed quantized ML to gain speed and save energy with minimal impact on accuracy
- Additional use cases: ultra-fast, on-edge, energy-efficient ML incl. fast triggering and classification of events, search engines and chatbots (FPGAs used in BING and SIRI), privacy-compliant inference (object classification without storing camera data)

Tech specs

- Pruning, quantization (binary, tertiary), compression and parallelization of models by ML experts
- hls4ml integrates with DL libraries: reads as input models trained with Keras/TF, PyTorch, scikit-learn, planned xgboost and outputs hls code; uses Xilinx HLS software (accessible to non-expert, engineers)
- Inference time: <10 microsec
- Comes with implementation of common network components (layers, activation functions, binary NN, …); example classification network: 16 inputs, 3 layers with 64/ 32/ 32 nodes (ReLU) and 5 outputs (Softmax)

Added value

CERN can help organizations design optimized neural networks for neuromorphic chips

- Adapt neural network code for edge inference on FPGA architecture with hls4ml
  → Reduced fixed-size development cost for edge inference, enhanced privacy by enabling transfer learning on data
- Design neural networks for fast and efficient inference through pruning, quantization, compression, parallelization
  → Energy-efficient, low-cost, reduced latency, ultra-fast inference
- Understand latest developments, define strategy for adoption and R&D
  → Advice and opinion on fast ML and neuromorphic chips by ML experts

Image credits: www.pexels.com, www.unsplash.com, CERN

ZENSEACT (Volvo Cars Company) teams up with CERN on fast machine learning using FPGAs.

Collaborative R&D

- General issue
- Jointly find solution
- Jointly develop solution

Collaborative R&D

- General issue
- Jointly find solution
- Jointly develop solution

CEVA and CERN joined R&D on neural network weight and activation compression algorithms aiming make them run more efficiently. Wireless comms & computer vision applications

Neural Network Learning and Optimization (NNLO)

Currently under development…a KT Funded project aiming to create a service for Industry (and our community)

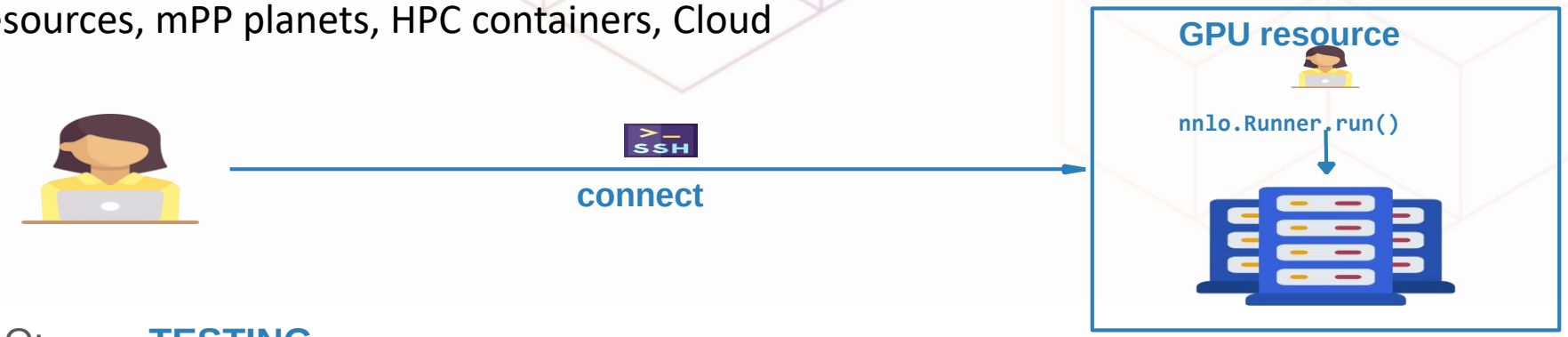
- **Software library** for distributed training and optimization.

Objectives:

- Unified distributed deep learning training workflow
  - ✓ Common interface for Tensorflow and Pytorch
  - ✓ Distribution strategy → Single GPU, → Single node with multiple GPUs, → Multiple nodes
  - ✓ Simplified distributed hyperparameter optimization
  - ✓ Unified interface for samplers and pruners (Bayesian optimization, Hyperband,...)
  - ✓ Pluggable optimization frameworks (Optuna, Keras tuner,...)

Deployment on diverse platforms

Local resources, mPP planets, HPC containers, Cloud



Status in NNLO: **TESTING**

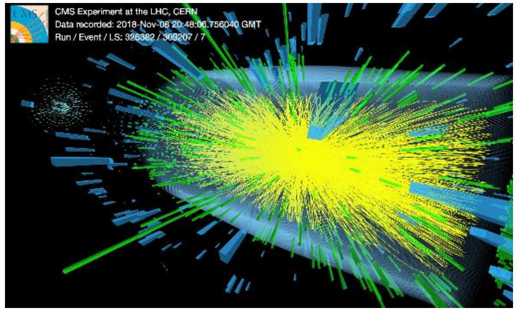
Container-based workflow

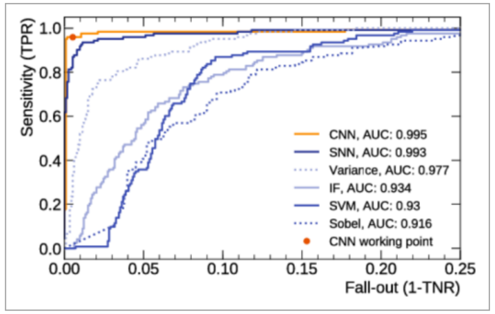
Export self-contained containers that can run on the target platform

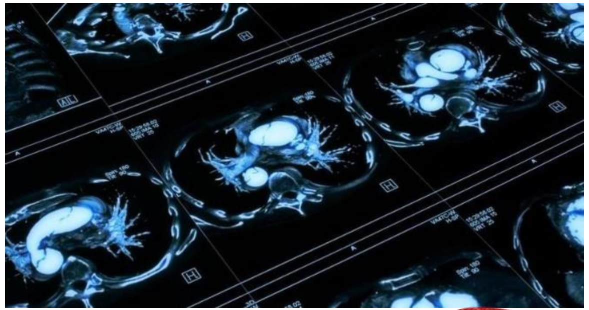
- Singularity for HPC, Docker for cloud (Kubernetes/Kubeflow)

Key competence: Big data classification and anomaly detection

CERN researches and operates and highly sensitive ML models for the detection of weak signals in very large datasets







What

CERN generates and stores large data volumes (1,000 observations/ sec. corresponding to 1 GB/ sec. and 10,000 TB/ year). A single physics analysis typically involves millions of observation examples to reach science-grade results requiring strong classification and anomaly detection algorithms.

For this, CERN researchers develop deep neural networks to effectively reject background noise from weak signals as well as a modular big data software framework (ROOT) for data handling, analysis, and visualization.

Apps

- An **institute for commodity risk** management teamed up with CERN to support regulators to detect **trading anomalies** from stock market data
- Knowledge Transfer supports the development **anomaly detection** and classification algorithms for medical image analysis to **diagnose cancer and Covid-19 pneumonia**
- Additional use cases: signal detection in large-scale, noisy, high-dimensional data, such as identifying cybersecurity and fraud attacks, detecting dangerous goods in logistics data, detecting energy consumption anomalies, pharma quality control

Tech specs

- Data: 2bn recorded events/ year with 100m dimensions per example
- FCN, AE, LSTM[1] for unsupervised anomaly detection: comparing latent space representation (AEs)/ prediction (LSTMs) with observed data
- CNNs for anomaly detection through image recognition: plotting (sensor) read-outs as image and training CNN to recognize anomalous images for e.g., data quality monitoring (reached AUC = 0.995)
- Weak signal detection where signal occurs at ratio of 1 in 10^3 to 1 in 10^6
- Example NN: parameters: 2.3m, epochs: 100, examples: 100k – 1m
- Boosted Decision Trees (ROOT, XGBoost) to improve data resolution

Added value

CERN can help organizations use its models and know-how in anomaly detection

- Adopt or design and train fit-for-purpose models to detect anomalies in big data (supervised, unsupervised, semi-supervised)
  → Weak signal detection incl. from unlabeled data, fast model development
- Optimize model interpretability and transparency along science-grade standards
  → Develop trust in models by reporting the right metrics along with results
- Use CERN data as testbed for development of anomaly detection models
  → Assure model quality by benchmarking with high-quality, homogeneous big data

FCN: fully connected network, AE: autoencoder, LSTM: long-short-term memory

Images: www.pexels.com, www.unsplash.com, CERN

Collaborative R&D

- General issue
- Jointly find solution
- Jointly develop solution

Collaboration with CORMEC and WUR to support national banks and regulators to detect trading anomalies in stock market

DE TELEGRAAF
DINSDAG 22 FEBRUARI 2022

Nederlandse universiteiten trekken op met CERN

DE FINANCIËLE TELEGRAAF T25

Met kerngeleerden op zoek naar beursfraude

door Dorinde Meuzelaar

AMSTERDAM • Beursfraude bestuderen met behulp van deeltjesfysica: dat is waar Wageningen Universiteit & Research en Maastricht University met onderzoekers aan werken. Op termijn hopen de onderzoekers misstanden zelfs te kunnen voorspellen.

Deeltjesfysica, een landbouwuniversiteit en de beursvloer: op het eerste oog liggen die niet in elkaars verlengde. Maar zo gek is de samenwerking niet, zegt hoogleraar marktkunde Joost Pennings. „In deeltjesfysica genereer je enorm veel data, en zoek je daarin naar de botsingen die niet perfect zijn." En hoewel het in de natuurkunde gaat over de bouwstenen van atomen in plaats van beurshandel, is er wel een overlap met financiële markten, waar in een duizelende tijd van een seconde heel veel transacties plaatsvinden. Door te zoeken naar 'anomalieën', of afwijkende patronen, kun je opsporen of het gaat om verboden handelspraktijken.

Spoofing

De onderzoekers bestuderen onder meer voorbeelden van 'spoofing' – een marktjargon die eerst orders plaatst maar die op het laatste moment weer annuleert. Zo lijkt het of er bijvoorbeeld veel interesse in een aandeel is, waardoor de prijs stijgt. Of juist anders om. De partij die aan 'spoof' doet, beïnvloedt dus tijdelijk de koers – om daarzelf aan te verdienen.

In het kort

- De universiteiten van Wageningen en Maastricht trekken samen met kerngeleerden op de beurs te bestuderen.
- Zo kunnen ze toezichthouders helpen bij het onderzoeken van misstanden, om de beurs veiliger te maken.
- De onderzoekers zien ook buiten de financiële markten toepassingen voor het onderzoek.

En dat gaat vaak razendsnel, zegt de Wageningse onderzoeker Marjolein Verhoest, die een spoofing-zaak van JPMorgan van 2008 tot 2016 bestudeerde. „In die zaak was de kortste interval van een spoof-order, dus tussen plaatsen en weer intrekken, 0,3 seconden." Het knipperen met je ogen duurt ongeveer even lang.

De universiteiten werken niet alleen samen met CERN, maar ook met verschillende toezichthouders zoals De Nederlandsche Bank (DNB) en de Autoriteit Financiële Markten (AFM). De AFM speurt zelf al in transactiedata, op zoek naar misstanden.

Pennings: „Wat wij doen is daarop een aanvulling. Zie het als een zaklamp. De AFM heeft al een heel sterke zaklamp, maar die van ons schijnt licht op andere zaken."

Door dit soort zaken beter op te sporen, moeten financiële markten beter worden. Pennings: „Er is een reden waarom dingen niet mogen. Je wil geen valse informatie in de markt, waar andere mensen weer te goeder trouw op handelen."

In de toekomst hoopt Pennings dat het zelfs te voorspellen is wanneer een overtreding plaatsvindt. „Vergelijk het met voorspellen van een tornado. Je voelt de luchtdruk en luchtvochtigheid veranderen, de temperatuur verandert; dat zijn allemaal vrij goede voorspellers of er een tornado aankomt of niet. Misschien kunnen we zo ook naar beurstransacties gaan kijken, dat we vooraf weten onder welke omstandigheden we bepaalde overtredingen zien. Al is dat voorlopig nog toekomstmuziek."

Verhulst ziet bovendien dat deze manier van grote hoeveelheden data bestuderen ook buiten de beurzen kan worden toegepast. „Banken kunnen zo gerichter zoeken in creditcard-transacties. We hebben

Overtredingen op de beurs gaan razendsnel: knipper met je ogen en het is al gebeurd.

FOTO GETTY IMAGES

'Voorspellen wanneer overtreding plaatsvindt'

COLUMN

JANNEKE WILLEMSE

Voorlopige aanslag

Sinds ik ondernemer ben, moet ik ieder jaar de belasting vooruitbetalen. Daar schrok ik het eerste jaar als zzp'er enorm van. Want in dat eerste jaar moest ik dubbel afrekenen: de voorlopige aanslag over het komende jaar, waarin ik nog gewoon in loondienst was. Na dubbel- en driedubbelchecken via Google en bij mensen die het konden weten en nadat ik erachter was gekomen dat dit kennelijk de normale gang van zaken was, heb ik de voor- en na-aanslagen allebei betaald. Natuurlijk, het was even slikken, maar ik kon het betalen van de buffer die ik lers tijdens mijn nieuwe ondernemede leven mee op te vangen.

Na die eerste keer raakte ik eraan gewend mijn belasting van tevoren al af te rekenen. In de rest van het jaar vul ik telkens de buffer op mijn spaarrekening weer aan, zodat ik ook voor veel pijn kan afrekenen. Ik maak de voorlopige aanslag altijd zo snel mogelijk over naar de Belastingdienst, zodat ik er 13 betalingskortingen kan opstrijken (grapje), maar vooral zodat ik me niet kan vergissen in het bedrag dat ik nog kan uitgeven. Ik had het voorlopige-aanslaggeheel mijn ondernemerschap, maar deel van mijn ondernemerschap, maar

Belastingdienst had mij juist moeten betalen

vorig jaar kreeg ik toch weer even last van belastingpijn. De belangrijkste oorzaak: onrecht. Ik moest namelijk €82 aan belastingrente betalen. Een bedrag van niks natuurlijk, maar het gevoél dat ik werd afgestraft omdat ik een beter jaar had dan ik had verwacht, was er niet minder om. Bovendien had ik zoals altijd een voorlopig bedrag vooraf overgemaakt, waarmee ik dus niet had kunnen beleggen en waarover ik rendement had misgelopen. Als iemand rente zou moeten betalen was het de Belastingdienst, aan mij!

Met dit gevoel van onrecht heb ik verder niets gedaan. Mijn belastingadviseur rond het rentevrerhaal volkomen normaal en het is me te veel moeite, dus heb ik het erbij laten zitten. Ook bij de vermogensrendementsheffing heb ik eigenlijk nooit vraagtekens gezet. Maar wat ben ik blij met de 60.000 mensen die dat samen met de Bond voor Belastingbetalers wel hebben gedaan. Dankzij hen ligt er nu een arrest van de Hoge Raad, waaruit blijkt dat belasting betalen over een verzonnen rendement niet eerlijk is. Dankzij hen moéten er nu een beter systeem bedacht worden door het kabinet. We zullen nog tot mei moeten wachten tot duidelijk wordt of ook niet-klagers gelet terugkrijgen, maar ik klaag sowieso niet. Dankzij het arrest kan ik voorlopig mijn buffer weer parkeren op de piek waar ik er altijd snel bij kan: de spaarrekening. Een hoog fictief rendement behalen hoeft niet. De buffer wordt weer gewoon voorlopige aanslag.

MONTEURS VOOR 1850 KM KABEL

maandag fors te te blijven investeren in grote verdeelstations, het schema ligt volgens bestuursvoorzitter Bogers op twee tot vier stuks per jaar. Dat is een verdrievoudiging ten opzichte van het oude ritme, maar volgens Bogers broodnodig om Ne-

Project Piet Heinplein Den Haag

EIGEN FOTO

derland stevige infrastructuur te blijven bieden die de overgang van energievormen moet steunen.

'Nog jaren nodig'

Vooral voor groei van het elektriciteitsnetwerk in de bebouwde gebieden loopt Stedin tegen krapte aan: het moet ruimte reserveren, maar de plekken zijn schaars. De hele vernieuwing van het netwerk gaat mede daardoor nog jaren duren, aldus Bogers in een toelichting.

„We krijgen steeds meer te maken met allerlei vormen van schaarste zoals beperkte financieringsmogelijkheden, maar ook tekorten aan materialen en geschikte vakmensen

beroep gedaan op het nieuwe kabinet, maar vraagt ook de provincies en gemeentes om de uitbreiding financieel te steunen. Ook andere netwerkbedrijven kloppen in Den Haag aan vanwege de enorme investeringen.

Meer ruimte reserveren is nodig in nieuwbouwwijken, nieuwe bedrijventerreinen en bij datacenters. Bogers meldde maandag bij de jaarcijfers ook dat het aantal nieuwbouwwoningen zonder aardgasaansluiting in 2021 is gestegen naar

Contract Research

- Use case and requirements by the company
- Code contributed to the OS project
- Development @CERN, benefit for HEP applications

ROCHE is using CernVM-FS for application and library distribution worldwide.
Contract Research for a Company in the financial services sector. Strong interest in this tech for fast reliable worldwide file distribution.

Consultancy/Service

- Specific issue
- Time of experts
- Time of facilities

Bundesdruckerei GmbH works with CERN on next generation ideas for identity management and cryptography and data handling

Collaborative R&D

- General issue
- Case study
- Jointly develop solution

ABB teams up with CERN to build a digital twin of our cooling and ventilation system in order to optimize enegy usage

Advise / 2ND Opinion / Tech Challenge from CERN Expert team

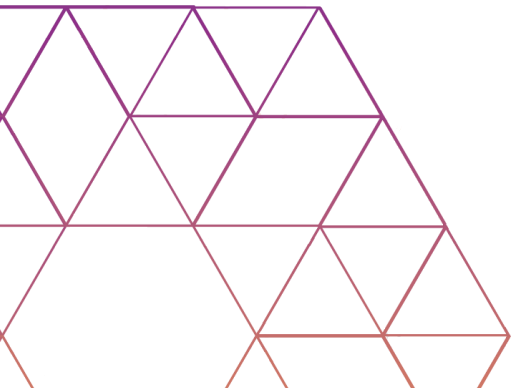
Collaborative R&D / Co-development agreements on specific topic of mutual interest

Challenge Based Innovation program with CERN Experts and/or universities to address specific issue

Using CERN labs / CERN openlab for joint R&D, prototyping, benchmarking, testing of software and equipment

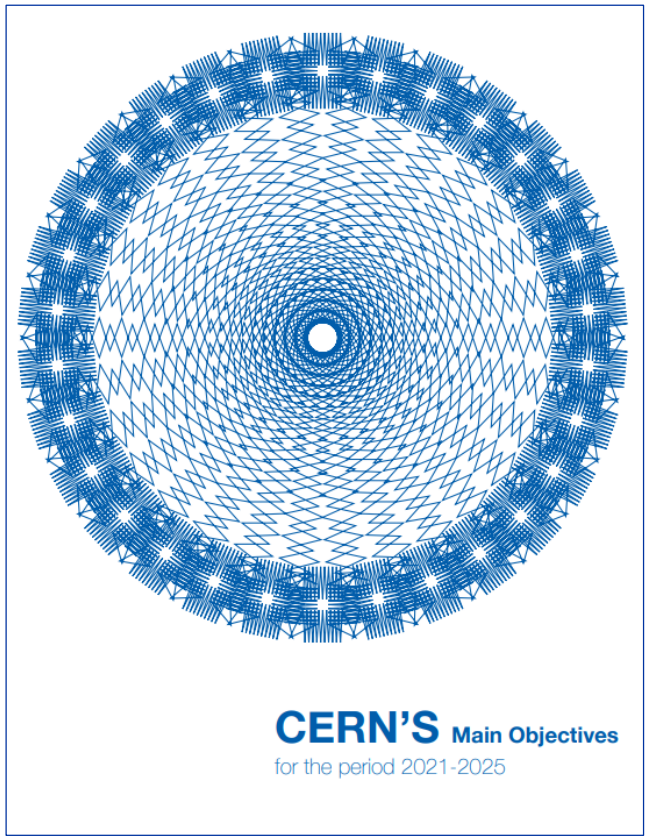
Licensing of CERN technology for commercial use / Support or training on using Open Source Hardware / Software

Facilitation of Knowledge Exchange by sponsoring PhD / Allocation of company resource at CERN / Use of Alumni Network



Environment: a clear priority for CERN

Three main development directions have been identified for environment and sustainability:



Minimise the Laboratory's impact on the environment by implementing CEPS (CERN Environmental Protection Steering) recommendations and defining a Green Procurement strategy

Identify and develop CERN's technologies that may contribute to mitigating the impact of society on the environment

CIPEA

Pursue actions and technologies aiming at energy saving and reuse, under the supervision of CERN's Energy Management Panel

CERN Technology Impact Fund

*A mechanism for seeking **donor funding** to support the further development of **CERN technologies** that have high potential to positively impact one or more of the 17 **United Nations Sustainable Development Goals (SDGs)***

- CERN personnel proposal for a high potential project that creates societal impact.
- Funding sought through the CERN & Society Foundation.
- Partnerships with external organisations in academia, the public sector and industry to maximise the chances of a successful technology transfer to society.

Further information about the SDGs: https://sdgs.un.org/goals

Nick Ziogas@cern.ch
cern.ch/kt