Integrating automation and optimisation into the CERN control system

A. Huschauer, N. Burchon, A. Calia, L. Felsberger, J.C. Garnier, R. Gorbonosov, M. Hostettler, F. Irannejad, D. Jacquet, V. Kain, N. Madysa, B. Rodriguez Mateos, K. Papastergiou, C. Petrone, M. Remta, M. Schenk, M. Sobieszek, G. Trad, F. Velotti, J. Wulff



Automation - key challenge for CERN accelerators

Long-term effort at CERN \rightarrow "Classical Automation" in the LHC



Automation - key challenge for CERN accelerators

Long-term effort at CERN \rightarrow "Classical Automation" in the LHC

optimisation algorithms, AI/ML

keynote

This is a kick-off discussion for a larger-scale effort in the making

Automating accelerators

- * Clear benefits for automating various processes will not go through the list of benefits
 - Automation means investment. Where do we have to automate and where should we automate?
- * CERN accelerator operation has progressively been more and more automated
- * Many of the essential building blocks are available or are in the making
- * Automation should become one of the complex goals for the years to come

Focus of today's meeting:

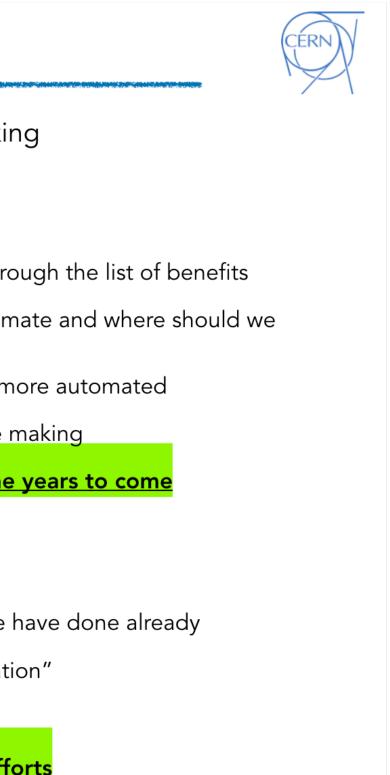
- * Show what is available in terms of building blocks, what people have done already
- * Try to define (or define how to define) "focus areas for automation"
- * Introduce the next step
- * → Establish community driven effort instead of individual efforts

IPP on Automation, Intro, 30/9/2022

SX Workshop - Wr. Neustadt, A. Huschauer & V. Kain, 14-Feb-2024



Revolution in the last years \rightarrow availability of controls building blocks for full automation,



Automation discussion becoming centre-stage:

Example: Introduction slide to Injectors Performance Panel (IPP) mini-workshop on automation on 30/9/2022



The waves of automation @ CERN

The **current** automation efforts are based on three threads

- Automation wave 1 (2006)
 - * reduce complexity through models in the accelerator control system (LSA)
 - * high level parameter control, procedures in software (sequencers), software interlock system, classic control algorithms in feedforward and feedback (SVD, COSE,...)



The waves of automation @ CERN

The **current** automation efforts are based on three threads

- Automation wave 1 (2006)
 - * reduce complexity through models in the accelerator control system (LSA)
 - * high level parameter control, procedures in software (sequencers), software interlock system, classic control algorithms in feedforward and feedback (SVD, COSE,...)
- Automation wave 2 (2018)
 - * provide clever solutions if models are not available. E.g. Learn them...
 - * game-changer: Python in the control room
 - * on-demand execution of **optimisers**, ML, etc.



The waves of automation @ CERN

The **current** automation efforts are based on three threads

- Automation wave 1 (2006)
 - * reduce complexity through models in the accelerator control system (LSA)
 - * high level parameter control, procedures in software (sequencers), software interlock system, classic control algorithms in feedforward and feedback (SVD, COSE,...)
- Automation wave 2 (2018)
 - * provide clever solutions if models are not available. E.g. Learn them...
 - * game-changer: Python in the control room
 - * on-demand execution of **optimisers**, ML, etc.
- Automation wave 3 (2021)
 - * Closing the loop...
 - Implementation of **frameworks** (Generic Optimisation Framework (GeOFF), Machine Learning Platform)
 - auto-launch correction, auto-resets of equipment, auto/online-analysis
 - enabling the operational implementation of auto-pilots



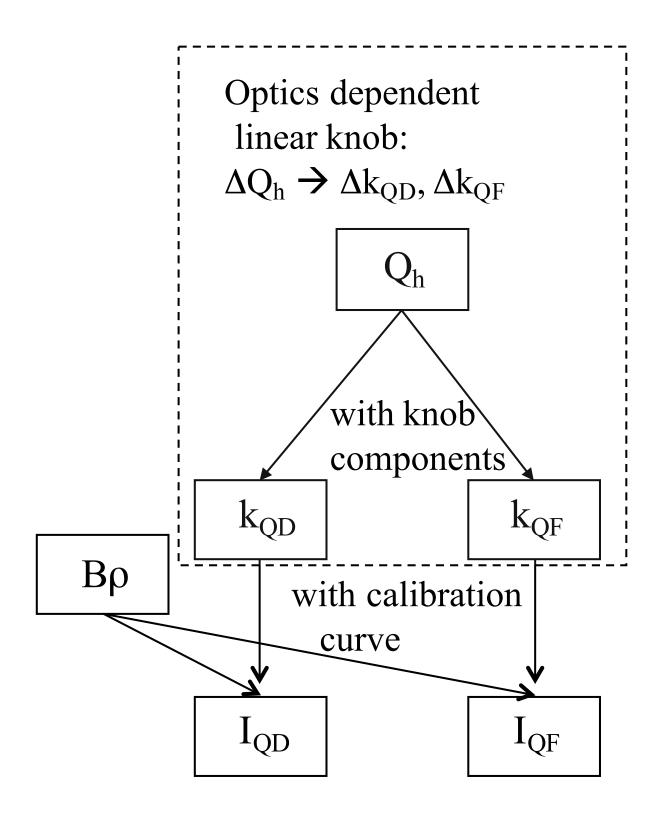
High Level Parameter Control and Generation

LHC Software Architecture (LSA)

- "only" work with normalised strengths k_n and high level physics parameters
 - * E.g. physics parameter Q_h defined on top of of k_{QD} and k_{QF}
- all parameters are functions of time in "cycle"
- The lower level field/current/... functions are automatically derived
 - * using centrally stored calibration curves, $B\rho$ /momentum, other parameters
 - $* \rightarrow$ key concept behind COSE

SX Workshop - Wr. Neustadt, A. Huschauer & V. Kain, 14-Feb-2024



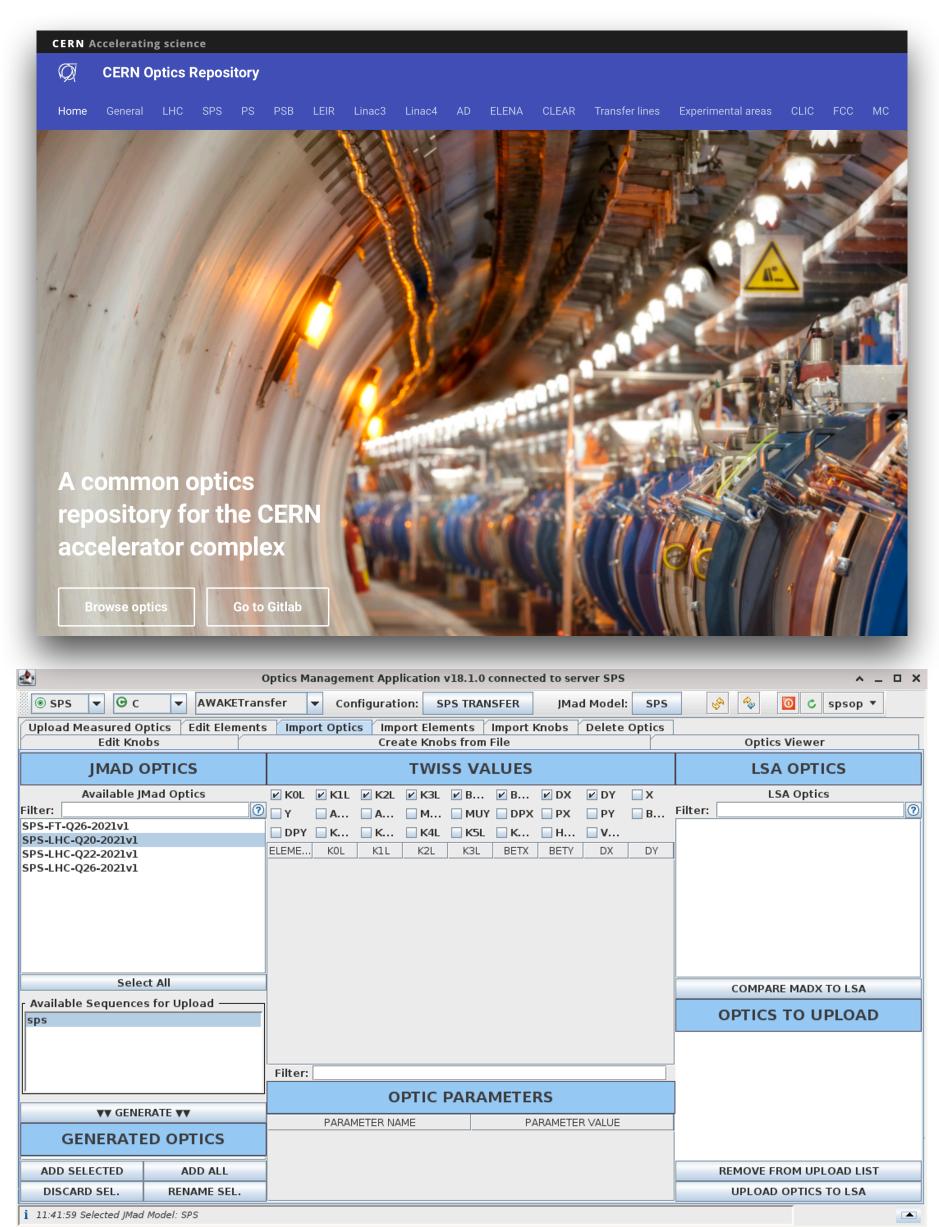


relations/rules/values in LSA DB...

High Level Parameter Control and Generation

Generation: generating parameters according to physics rules

- Parameter/settings generation based on MAD-X optics model rather than empirical definition
- Optics upload from gitlab-based <u>CERN optics</u> <u>repository</u>
 - * Efficient upload and modification of optics (even in the control room)





High Level Parameter Control and Generation

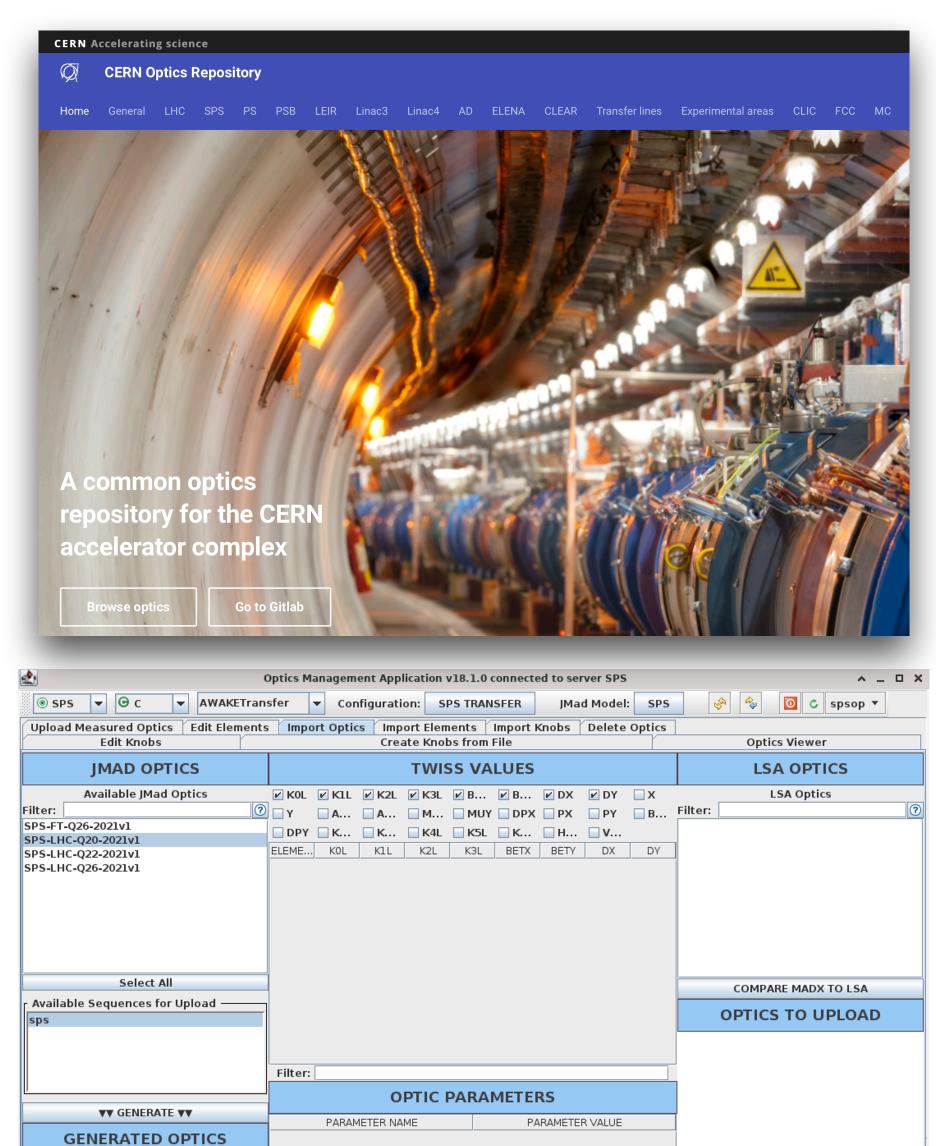
Generation: generating parameters according to physics rules

- Parameter/settings generation based on MAD-X optics model rather than empirical definition
- Optics upload from gitlab-based <u>CERN optics</u> repository
 - * Efficient upload and modification of optics (even in the control room)

HL parameter control is natural operating principle of the LHC and the SPS

Large effort in recent years to roll out in the injectors

SX Workshop - Wr. Neustadt, A. Huschauer & V. Kain, 14-Feb-2024



i 11:41:59 Selected JMad Model: SPS

ADD SELECTED

DISCARD SEL.

ADD ALL

RENAME SEL

_	_	_	-

REMOVE FROM UPLOAD LIST

UPLOAD OPTICS TO LSA

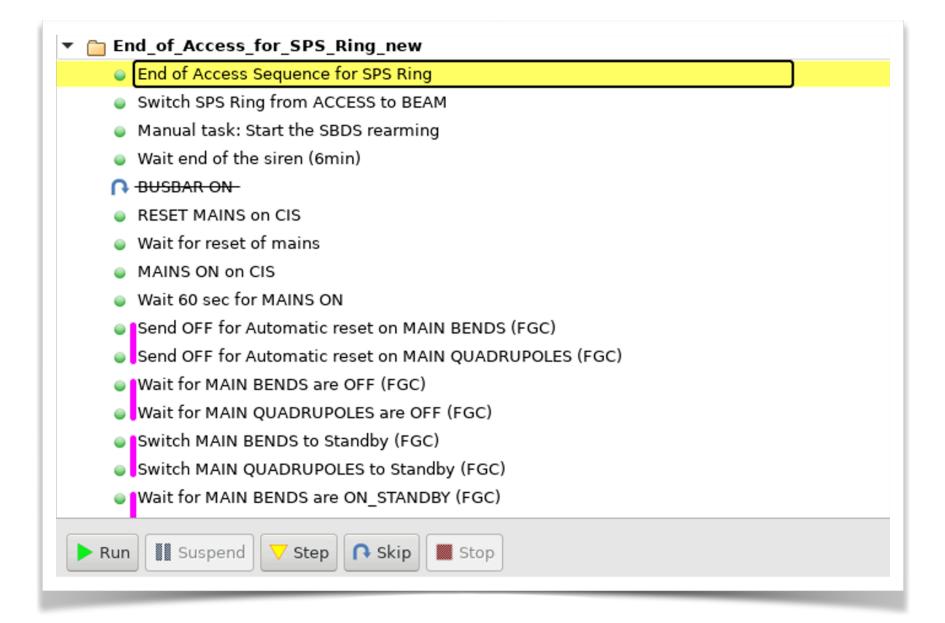


Sequencer

- = transforming procedures into code \rightarrow reusable sequences, sequence tasks,...
 - Benefit: clearly defined and hence reproducible sequence of actions
 Actions
 - Example: give and end access to SPS ring

Access_for_SPS_Ring_new	
Access Sequence for SPS Ring	
Manual task: Switch OFF Operator Button	
Mains CYCLING to ON_STANDBY (FGC)	
BHZ377 to ON_STANDBY	
BHZ378 to ON_STANDBY	
Wait for MAIN BENDS are ON_STANDBY (FGC)	
Wait for MAIN QUADRUPOLES are ON_STANDBY (FGC)	
Note: Wait for BHZ377 to be ON_STANDBY (FGC)-	
N Wait for BHZ378 to be ON_STANDBY (FGC)	
Switch Off MAIN BENDS (FGC)	
Switch Off MAIN QUADRUPOLES (FGC)	
Switch off Sextupoles and Octupoles (FGC)	
Switch off MAL1001M MBIV1003M (FGC)	
Switch OFF BHZ377	
Switch OFF BHZ378	
Wait for MAIN BENDS are OFF or FAULT_OFF (FGC)	
Wait for MAIN QUADRUPOLES are OFF or FAULT_OFF (FGC)	
Wait for Sextupoles and Octupoles are OFF or FAULT_OFF (FGC)	
Wait for MAL1001M MBIV1003M are OFF (FGC)	
Wait for BUZ377 to be OFF or FAULT OFF	
▶ Run III Suspend V Step Skip III Stop	

The entire LHC cycle is orchestrated by the LHC sequencer (GUI-managed)



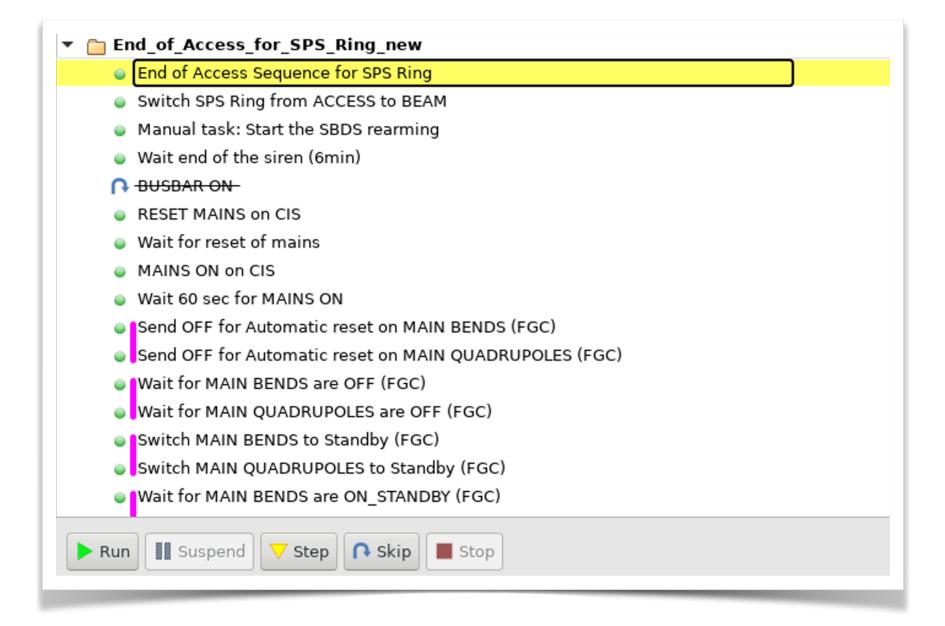


Sequencer

- = transforming procedures into code \rightarrow reusable sequences, sequence tasks,...
 - Benefit: clearly defined and hence reproducible sequence of actions
 Actions
 - Example: give and end access to SPS ring

Access_for_SPS_Ring_new	
Access Sequence for SPS Ring	
Manual task: Switch OFF Operator Button	
Mains CYCLING to ON_STANDBY (FGC)	
BHZ377 to ON_STANDBY	
BHZ378 to ON_STANDBY	
Wait for MAIN BENDS are ON_STANDBY (FGC)	
Wait for MAIN QUADRUPOLES are ON_STANDBY (FGC)	
N Wait for BHZ377 to be ON_STANDBY (FGC)	
N Wait for BHZ378 to be ON_STANDBY (FGC)	
Switch Off MAIN BENDS (FGC)	
Switch Off MAIN QUADRUPOLES (FGC)	
Switch off Sextupoles and Octupoles (FGC)	
Switch off MAL1001M MBIV1003M (FGC)	
Switch OFF BHZ377	
Switch OFF BHZ378	
Wait for MAIN BENDS are OFF or FAULT_OFF (FGC)	
Wait for MAIN QUADRUPOLES are OFF or FAULT_OFF (FGC)	
Wait for Sextupoles and Octupoles are OFF or FAULT_OFF (FGC)	
Wait for MAL1001M MBIV1003M are OFF (FGC)	
Wait for BUZ377 to be OFF or FAULT OFF	
▶ Run III Suspend V Step Skip III Stop	

The entire LHC cycle is orchestrated by the LHC sequencer (GUI-managed) key for auto-recovery, auto-measurement and auto-correction. SX Workshop - Wr. Neustadt, A. Huschauer & V. Kain, 14-Feb-2024



Next: sequences/sequence tasks to be **launched programmatically** (no GUI needed)



...with all the devOps one can dream of.

Python fully integrated with control system:

Equipment access: pyjapc, pyda
 * next: event building



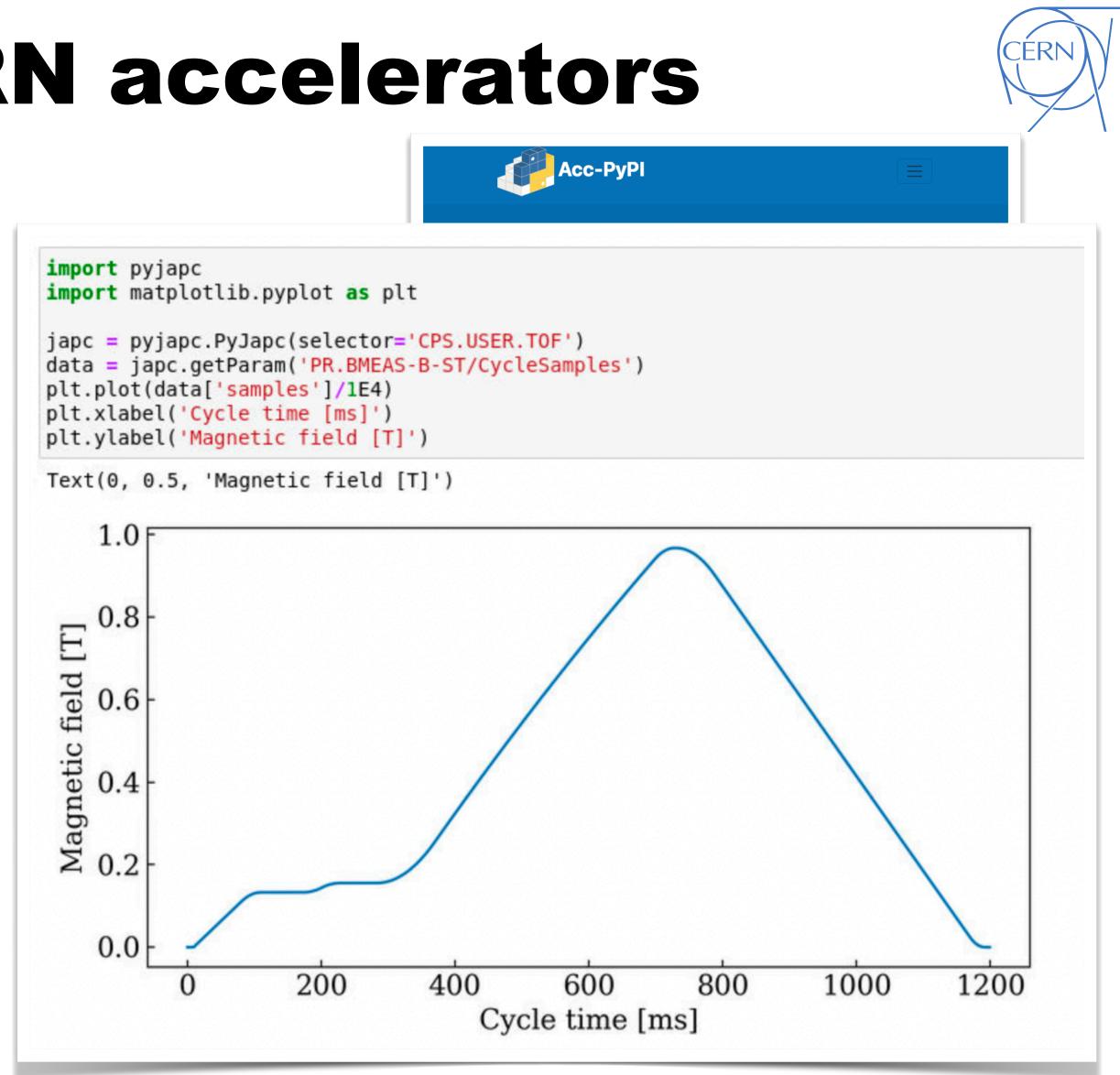




...with all the devOps one can dream of.

Python fully integrated with control system:

Equipment access: pyjapc, pyda * next: event building



...with all the devOps one can dream of.

Python fully integrated with control system:

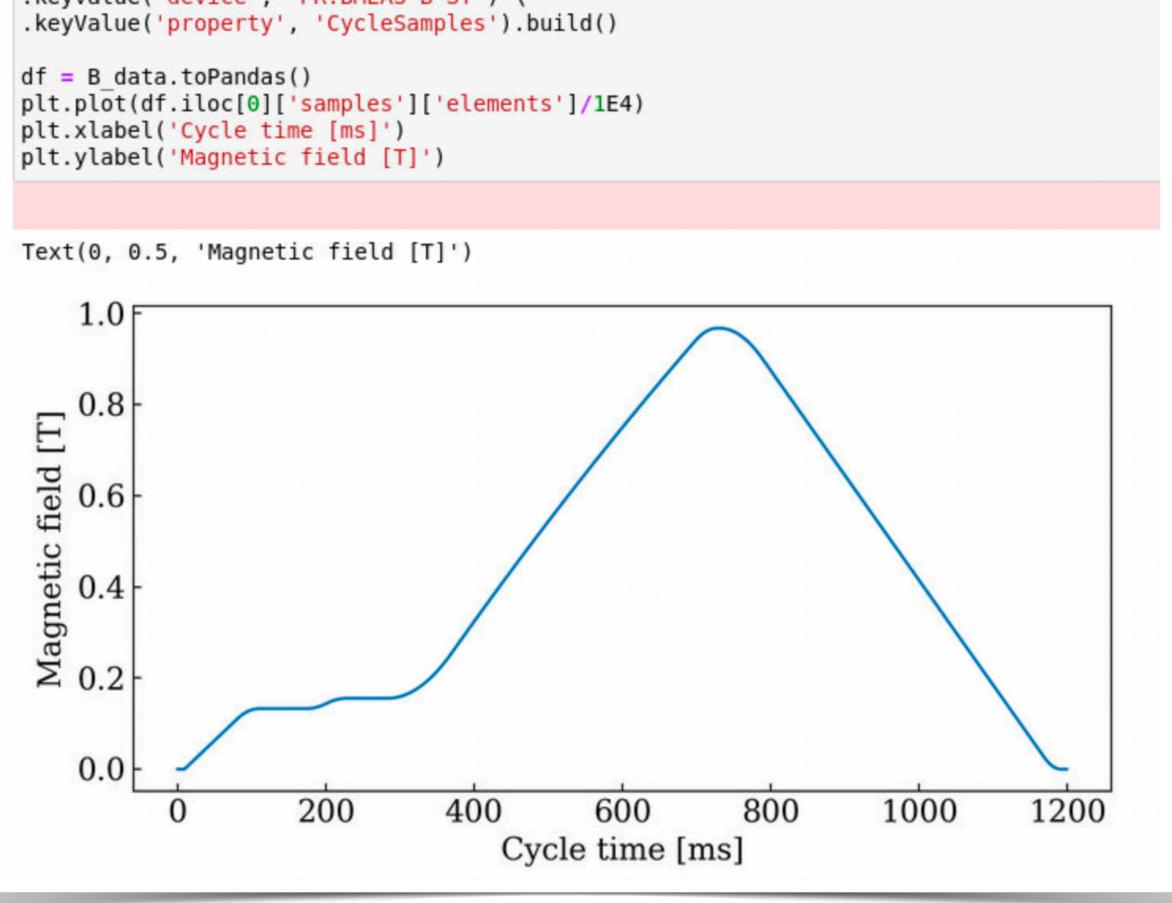
- Equipment access: pyjapc, pyda * next: event building
- access to settings (pjLSA) and archived data (NXCALS)

SX Workshop - Wr. Neustadt, A. Huschauer & V. Kain, 14-Feb-2024



Acc-PvPI

```
ts1 = datetime(2024, 2, 13, 8)
ts2 = datetime(2024, 2, 13, 8, 1)
B data = DataQuery.builder(spark).byEntities()\
.system('CMW').startTime(ts1).endTime(ts2)\
.entity()\
.keyValue('device', 'PR.BMEAS-B-ST') \
.keyValue('property', 'CycleSamples').build()
df = B data.toPandas()
plt.plot(df.iloc[0]['samples']['elements']/1E4)
plt.xlabel('Cycle time [ms]')
plt.ylabel('Magnetic field [T]')
```



...with all the devOps one can dream of.

Python fully integrated with control system:

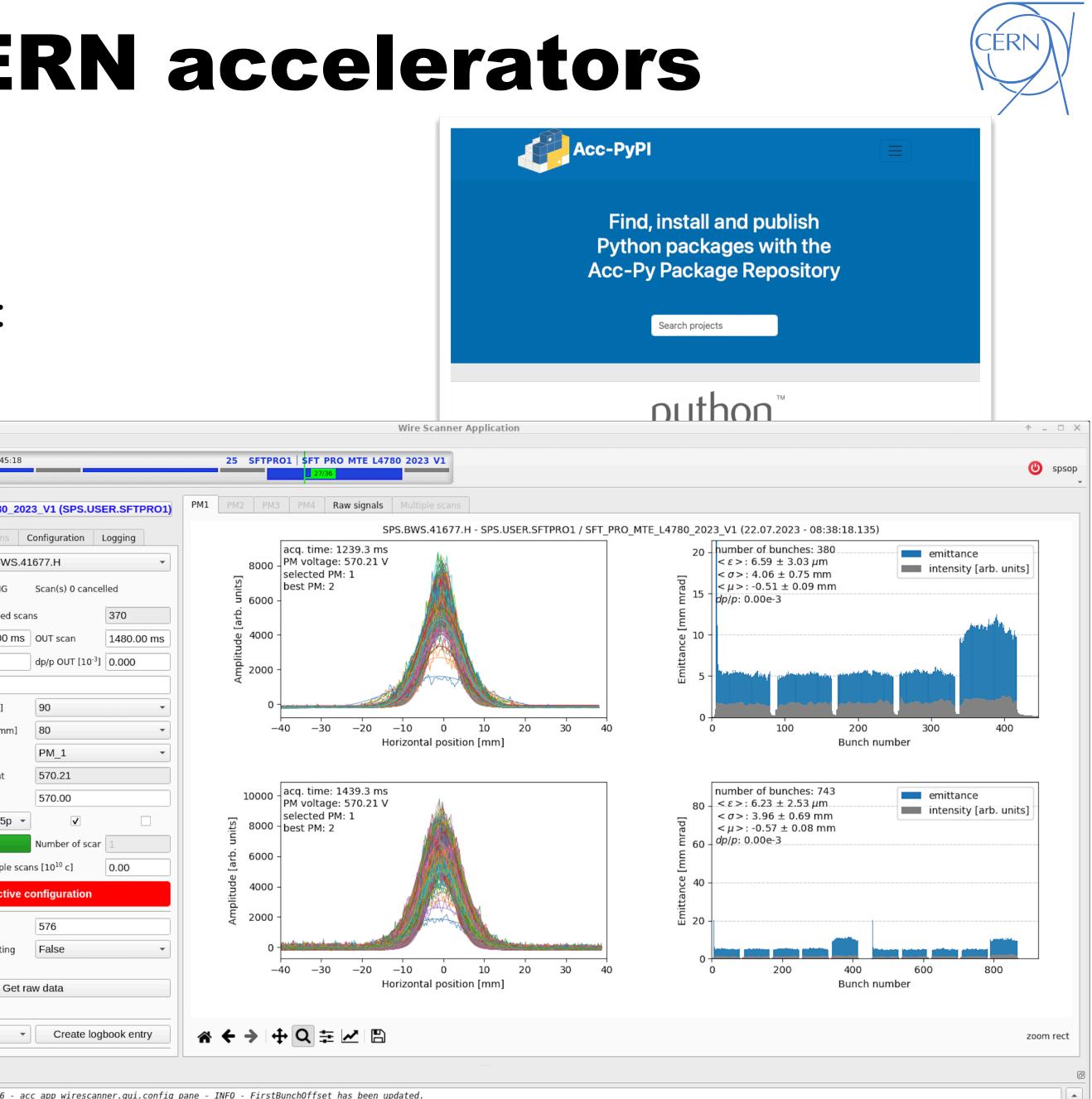
- Equipment access: pyjapc, pyda * next: event building
- access to settings (pjLSA) and archived data (NXCALS)
- GUI framework using PyQt and custom widgets
- and more...

SPS 2023	-07-22 08:45:18
SFT_PRO_M	ITE_L4780_202
Settings M	ultiple scans
Device	SPS.BWS.4
Status	WARNING
Total number	of performed sca
IN scan	1280.00 ms
dp/p IN [10 ⁻³]	0.000
Bunch selectio	or 1-924;
Scan window v	width [mm]
Analysis windo	ow width [mm]
PM selection	
PM voltage me	easurement
PM voltage set	tting
Fit function	Gauss5p 🔻
Laun	ch scan
Intensity cutof	f for multiple sca
	No active o
First bunch off	fset
Phase measur	ement setting
	Get ra

WS.41677.H

90

80



2023-07-22 08:44:51,136 - acc_app_wirescanner.gui.config_pane - INFO - FirstBunchOffset has been updated

...with all the devOps one can dream of.

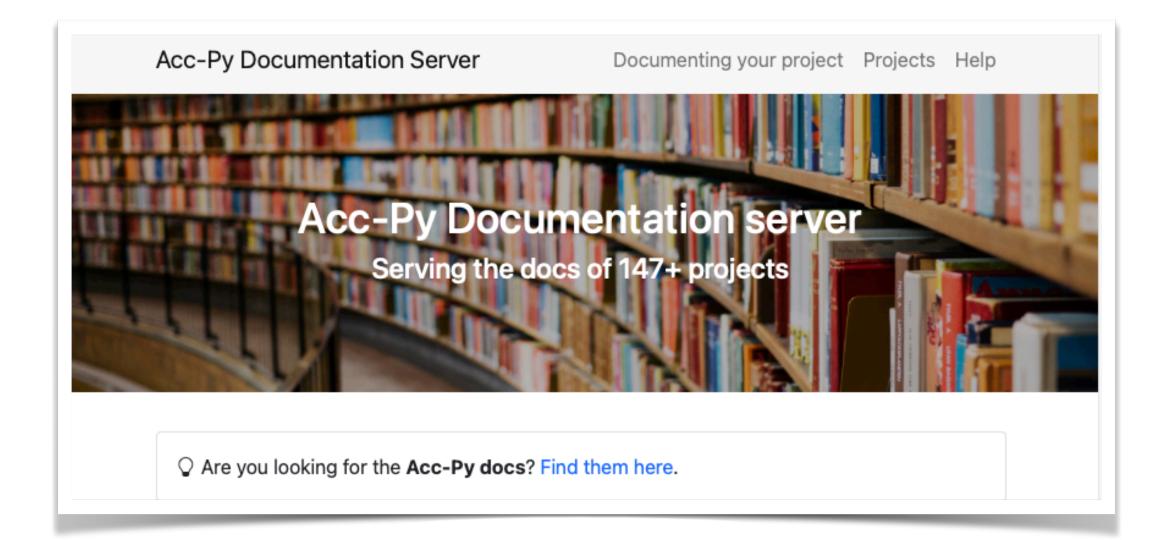
Python fully integrated with control system:

- Equipment access: pyjapc, pyda
 * next: event building
- access to settings (pjLSA) and archived data (NXCALS)
- GUI framework using PyQt and custom widgets
- and more...

Also: Acc-Py package indexing and central deployment location









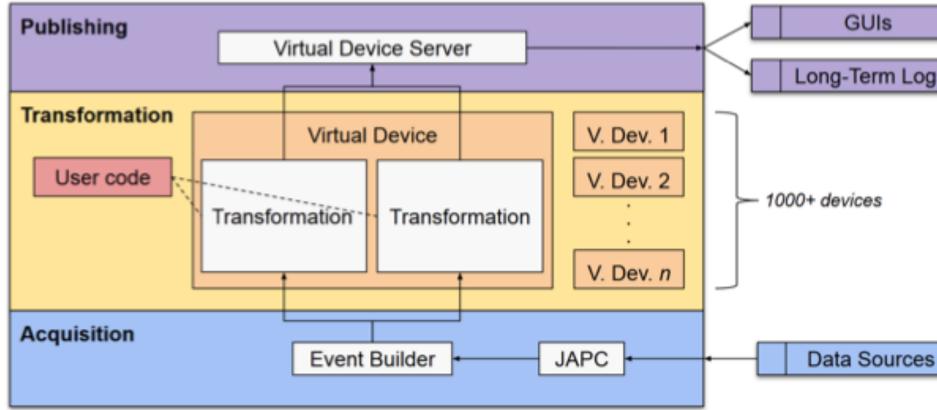
UCAP

Unified Controls Acquisition and Processing ("Virtual Device Service")

- servers on-the-fly in JAVA or Python for online data transformations



Plug&Play framework for online analysis (UCAP transformations), i.e. continuously running the background and triggered by "events"



g	in	
Uls		
erm	Logs	

UCAP

Unified Controls Acquisition and Processing ("Virtual Device Service")

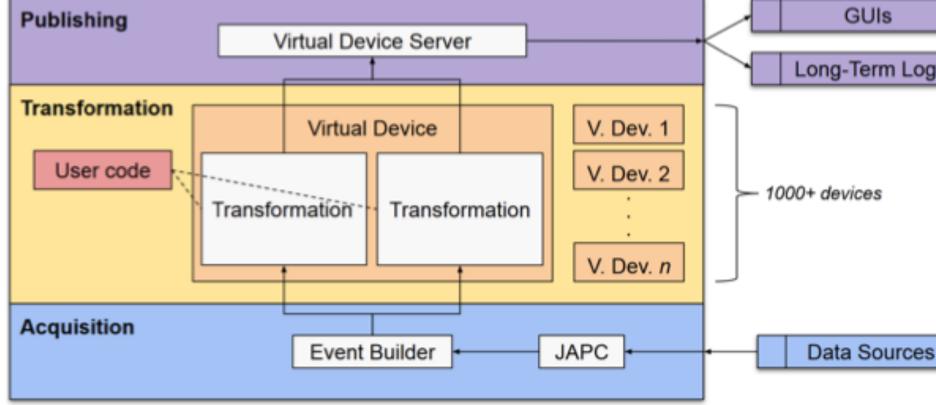
- servers on-the-fly in JAVA or Python for **online data transformations**

Provides:

- tools for development, test and deployment of "transformations" (+actors)
- centrally managed infrastructure to run transformations (GPUs available)



Plug&Play framework for online analysis (UCAP transformations), i.e. continuously running the background and triggered by "events"



g	in	
Uls		
erm	Logs	

UCAP

Unified Controls Acquisition and Processing ("Virtual Device Service")

- servers on-the-fly in JAVA or Python for **online data transformations**

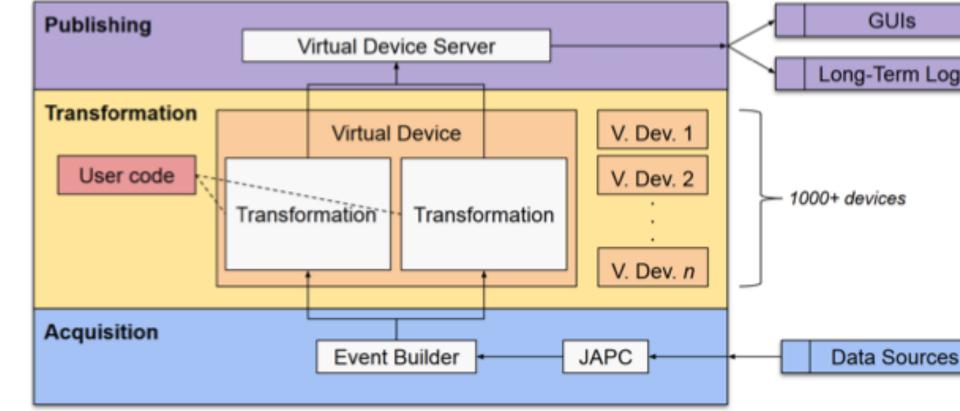
Provides:

- tools for development, test and deployment of "transformations" (+actors)
- centrally managed infrastructure to run transformations (GPUs available)
- Key features:
 - aggregating data: event building (e.g. buffered data,...)
 - software interlock systems, ...

SX Workshop - Wr. Neustadt, A. Huschauer & V. Kain, 14-Feb-2024



Plug&Play framework for online analysis (UCAP transformations), i.e. continuously running the background and triggered by "events"



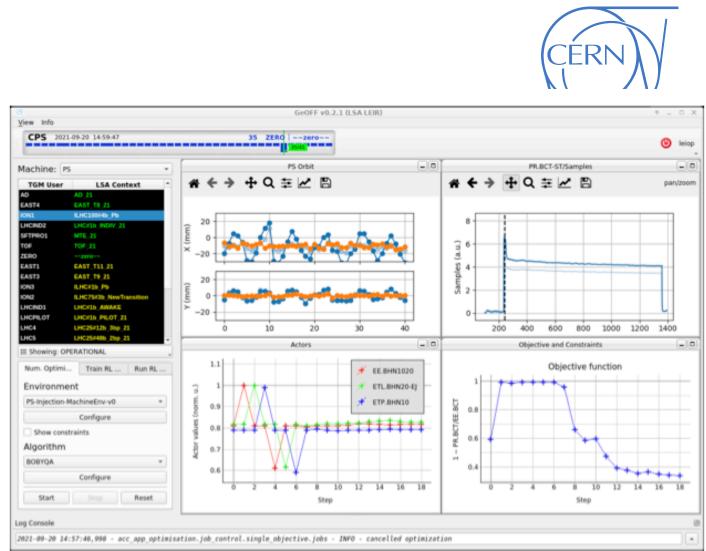
"transformation" result republished as online input for applications, controllers, logging,

g	in	
Uls		
erm	Logs	

Optimisation Infrastructure

Generic Optimisation Framework **GeOFF**

- **Replace manual grid scans** in the control room with **optimisation algorithms**
- GeOFF = flexible parameter optimisation (user-defined problem descriptions)
- To date: > 20 parameter optimisation problems automated across complex



Optimisation Infrastructure

Generic Optimisation Framework **GeOFF**

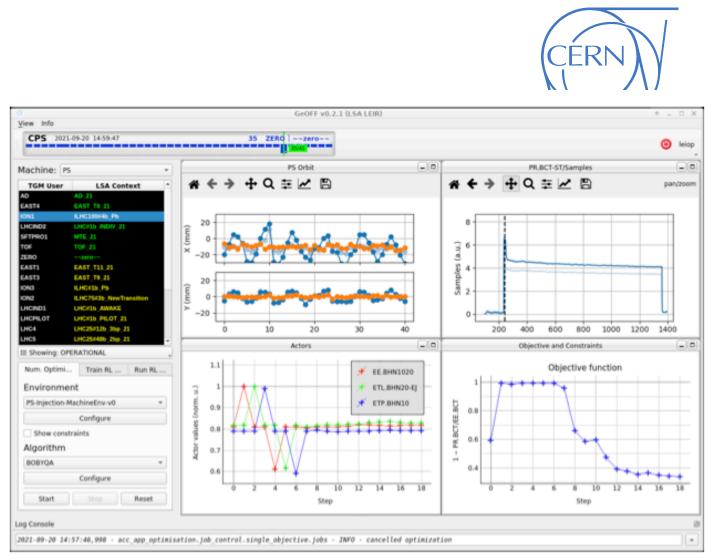
- Replace manual grid scans in the control room with optimisation algorithms
- GeOFF = flexible parameter optimisation (user-defined problem descriptions)
- To date: > 20 parameter optimisation problems automated across complex

Optimisation framework for auto-pilots

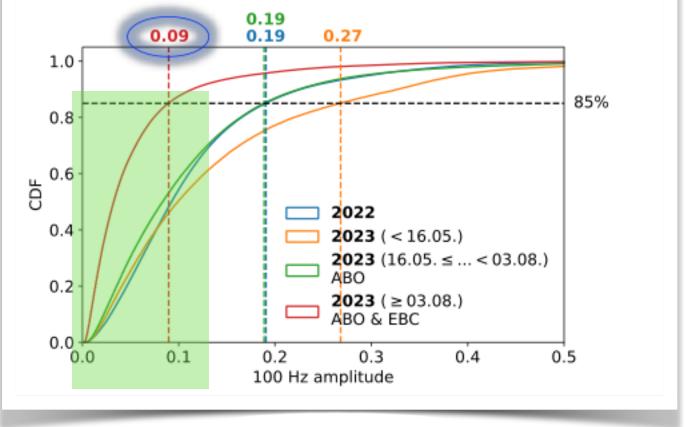


- GeOFF on UCAP \rightarrow acc-geoff4ucap released in summer 2023.
- **Operational**:
 - $* n \times 50$ Hz control with Adaptive Bayesian Optimisation (ABO) for North Area spill with GPUs on UCAP (see Francesco's talk)

SX Workshop - Wr. Neustadt, A. Huschauer & V. Kain, 14-Feb-2024



100 Hz content of NA spill with ABO and EBC





Optimisation Infrastructure

Generic Optimisation Framework **GeOFF**

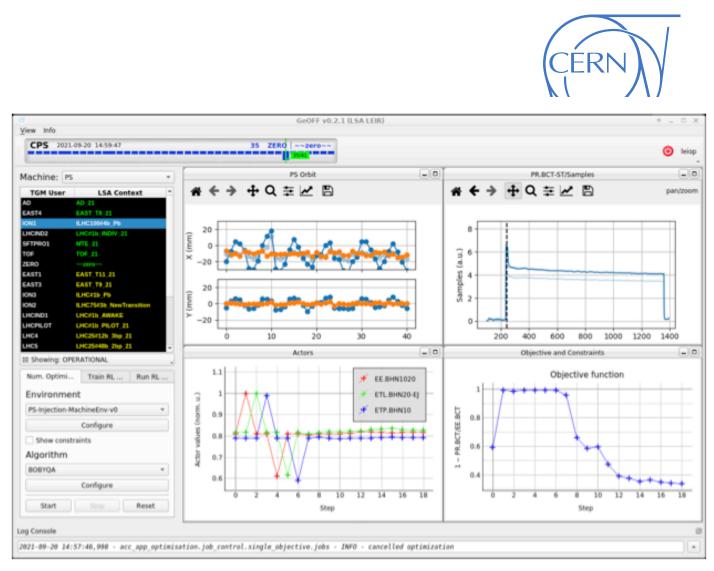
- Replace manual grid scans in the control room with optimisation algorithms
- GeOFF = flexible parameter optimisation (user-defined problem descriptions)
- To date: > 20 parameter optimisation problems automated across complex

Optimisation framework for auto-pilots

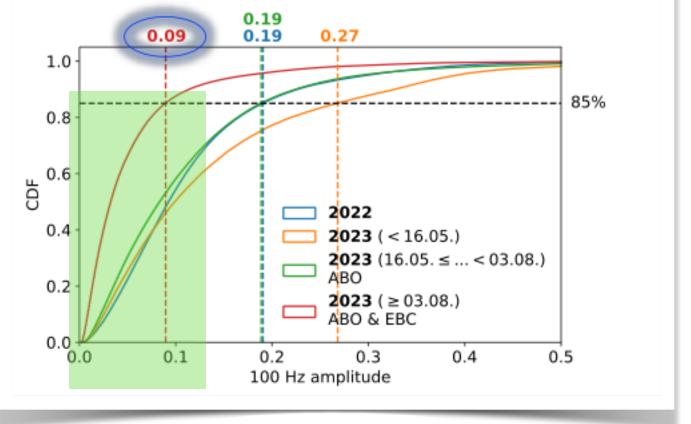


- GeOFF on UCAP \rightarrow acc-geoff4ucap released in summer 2023.
- **Operational**:
 - $* n \times 50$ Hz control with Adaptive Bayesian Optimisation (ABO) for North Area spill with GPUs on UCAP (see Francesco's talk)
- To come in **2024**:
 - * automated transfer line trajectory steering
 - * Multi-turn extraction (MTE) efficiency **drift stabilisation**

SX Workshop - Wr. Neustadt, A. Huschauer & V. Kain, 14-Feb



100 Hz content of NA spill with ABO and EBC



5-turn MTE spill: equal intensity sharing desired





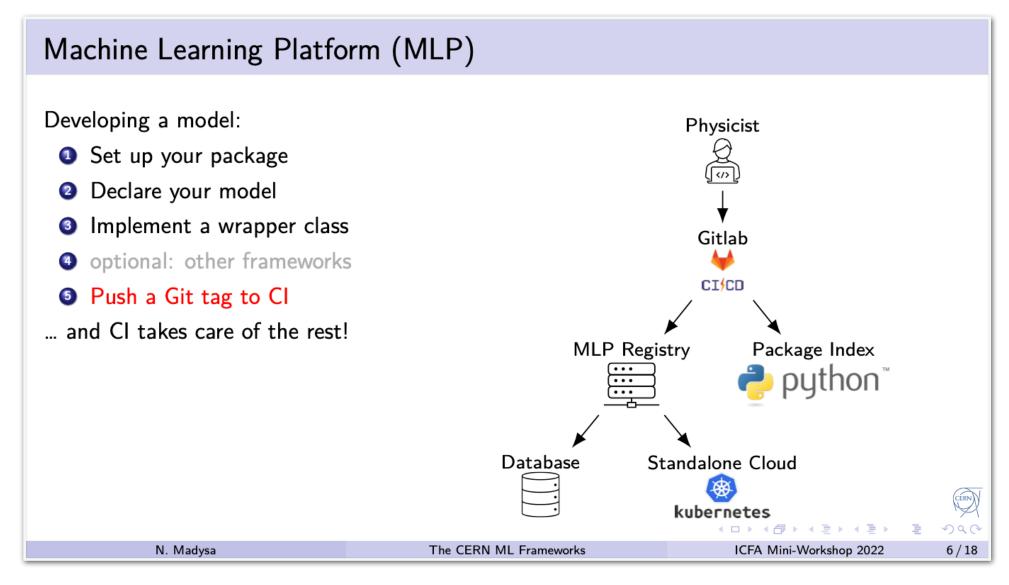




Machine Learning Platform (MLP)

different languages.

Stand-alone models hosted in the CERN cloud, ready for inference.



Successfully used for deploying AccGPT • challenge of very large Llama 2 model

SX Workshop - Wr. Neustadt, A. Huschauer & V. Kain, 14-Feb-2024



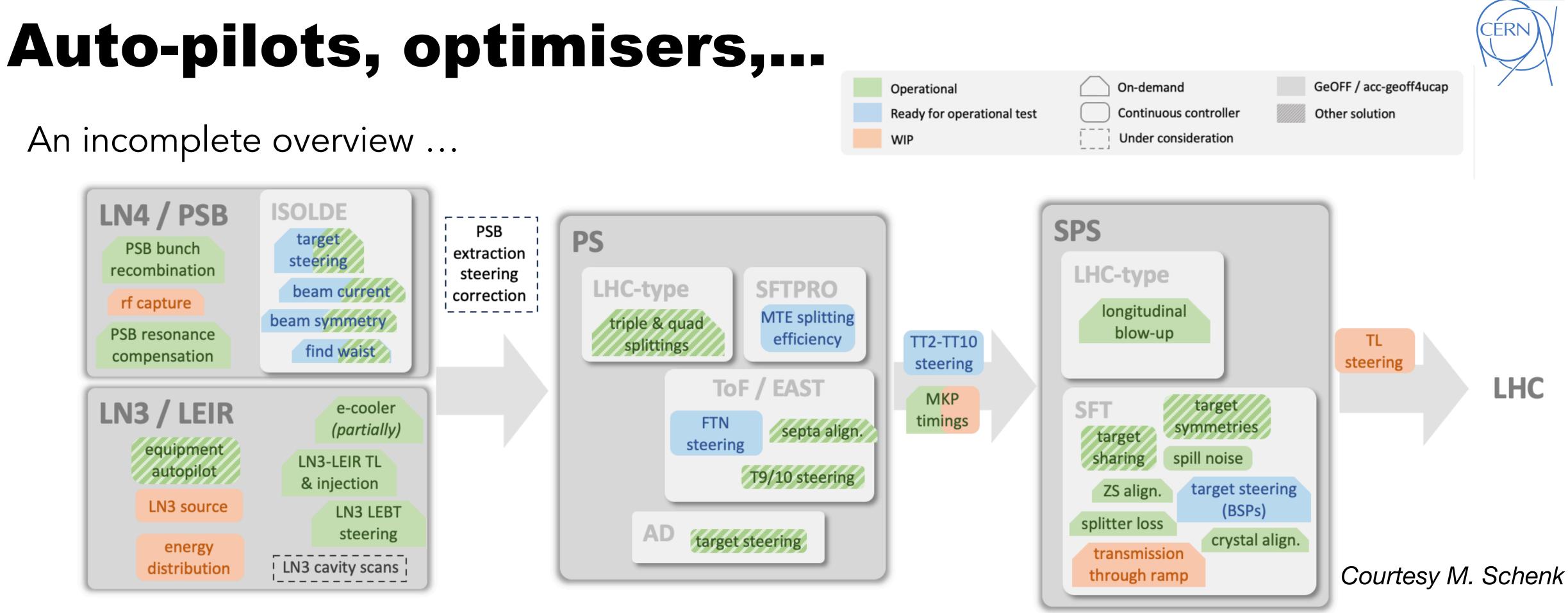
Storing and sharing trained ML models (e.g. ANNs) between users and applications of

```
Machine Learning Platform (MLP)
```

Connecting a Python application to a stand-alone model in the cloud:

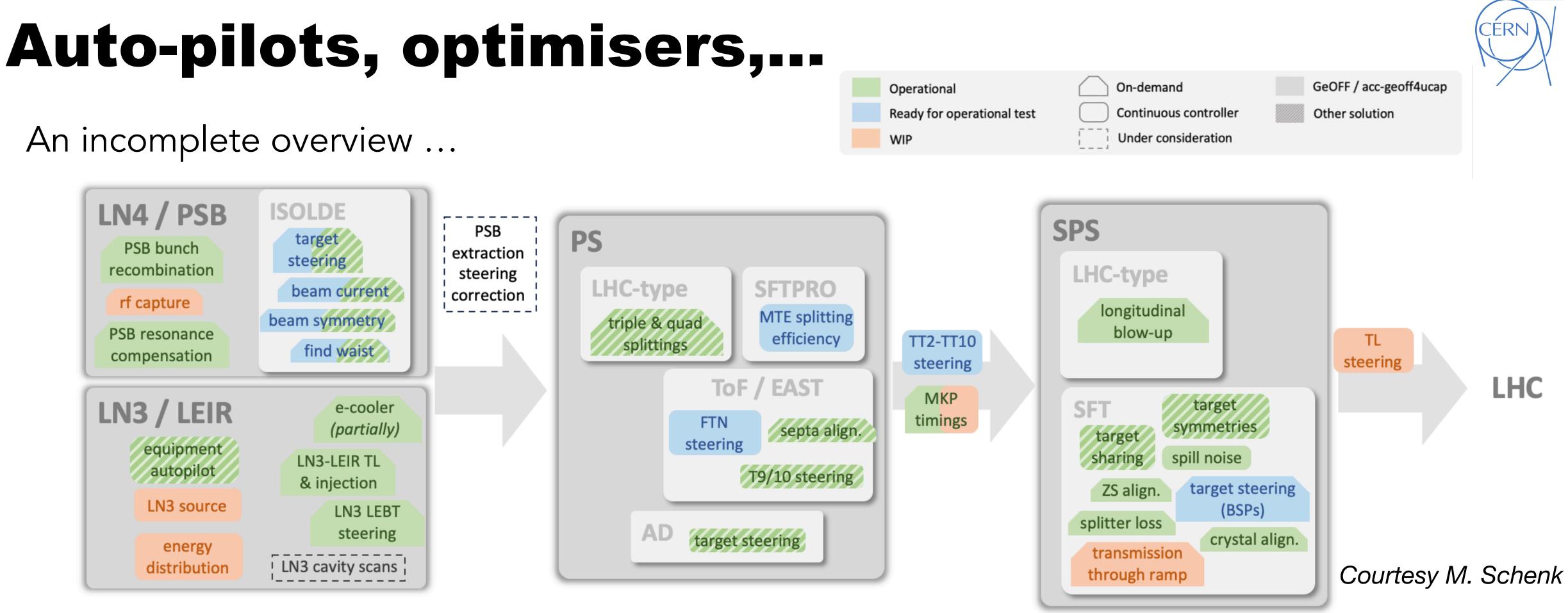
```
# operational_application.py
<sup>2</sup> from mlp_client import AUTO, Client, Profile
  from mlp_model_api import INPUTS, OUTPUTS
   client = Client(Profile.PRO)
   model = client.create_standalone_model(
       "my_model_package:MyModel",
       name="my_model.default",
8
       version=AUTO,
10
     Communicates via HTTP!
   response = model.predict({INPUTS: get_inputs()})
  show_results(response[OUTPUTS])
                                                          ▲□▶▲圖▶▲≣▶▲≣▶ ≣ めへで
```

An incomplete overview ...



Status 2023: multiple auto-pilots / optimisers used operationally and by experts during commissioning

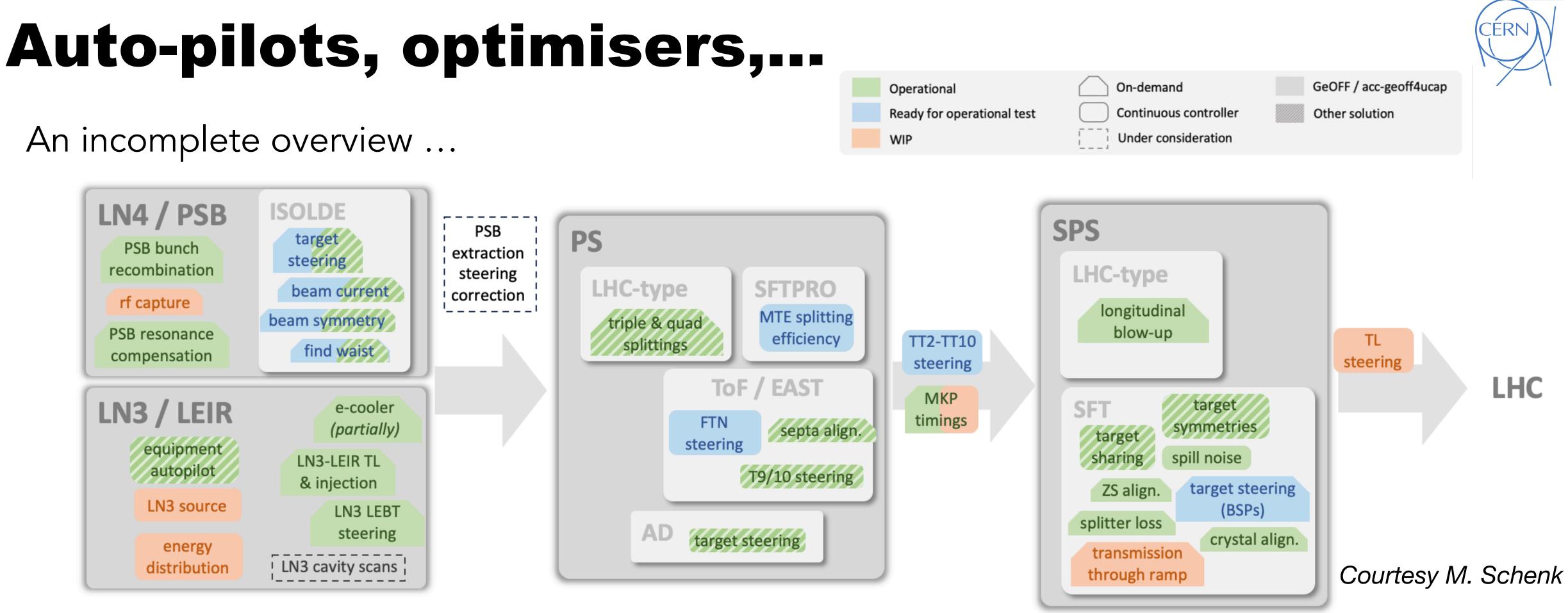
An incomplete overview ...



Status 2023: multiple auto-pilots / optimisers used operationally and by experts during commissioning

Trends 2024: move on-demand to acc-geoff4ucap and implement some new auto-pilots

An incomplete overview ...



Status 2023: multiple auto-pilots / optimisers used operationally and by experts during commissioning

Trends 2024: move on-demand to acc-geoff4ucap and implement some new auto-pilots

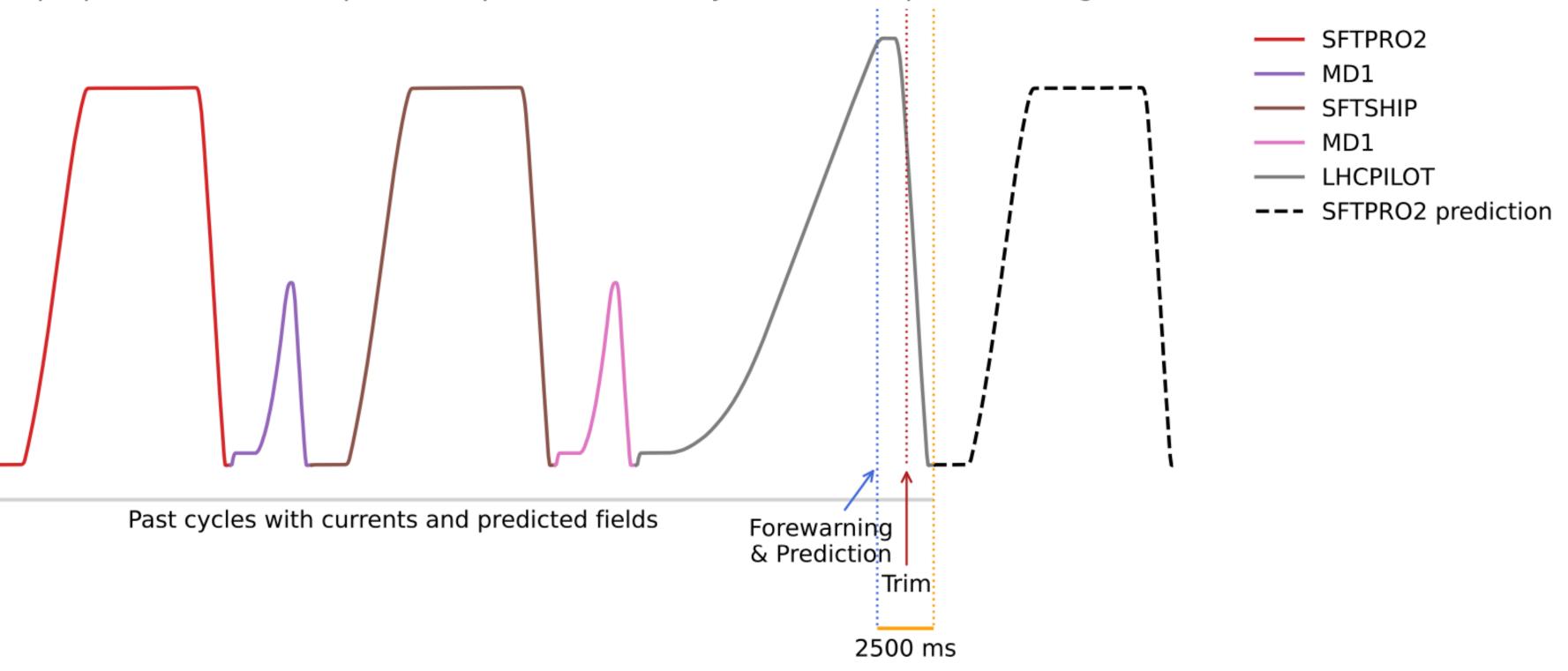
- Goal until end of run 3: automation of all typical optimisation and continuous control problems

Predicting magnetic hysteresis and eddy current effects

Training NN to predict future field and set correction $\hat{B}_{t+1} = NN(I_{t,t-1,...}, \hat{B}_{t,t-1,...})$

- Cycling the SPS is very expensive —> potentially game-changing
- feedforward correction triggered before every cycle

The proposed SPS main dipole field prediction and hysteresis compensation algorithm



SX Workshop - Wr. Neustadt, A. Huschauer & V. Kain, 14-Feb-2024





Courtesy A. Lu

Predicting magnetic hysteresis and eddy current effects

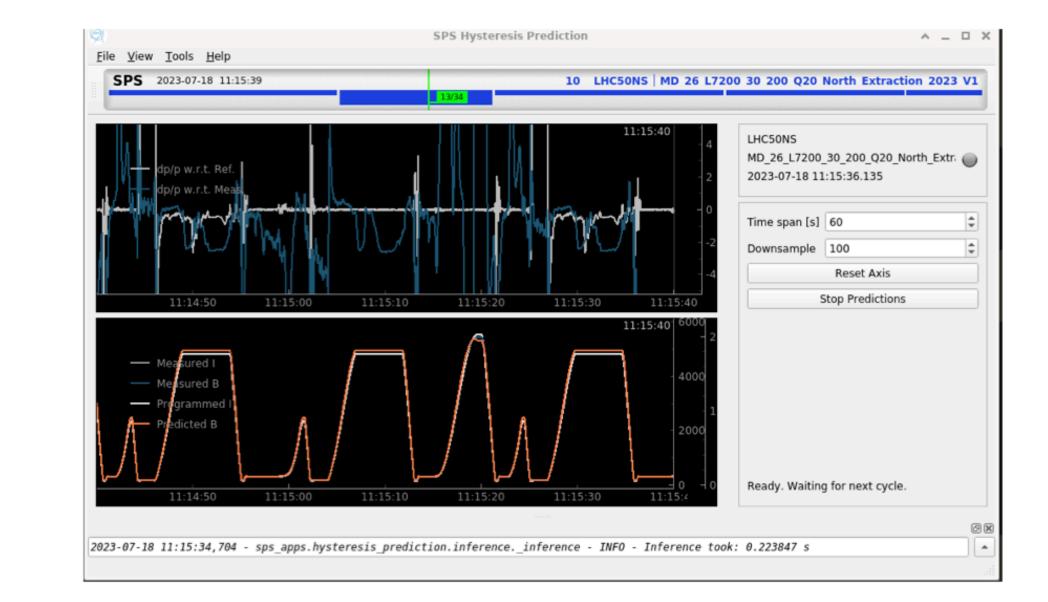
Training NN to predict future field and set correction $\hat{B}_{t+1} = NN(I_{t,t-1,...}, \hat{B}_{t,t-1,...})$

- Cycling the SPS is very expensive —> potentially game-changing
- feedforward correction triggered before every cycle

First operational experience

- first results using physics-inspired NN PhyLSTM $\ddot{B} + g = \Gamma i(t)$, $g = \left[b(B, \dot{B}) + r(B, \dot{B}, B(\tau))\right]/a$
- accuracy not yet sufficient
- Next:
 - Investigate usage of transformers
 - Collect more data of different super cycle configurations

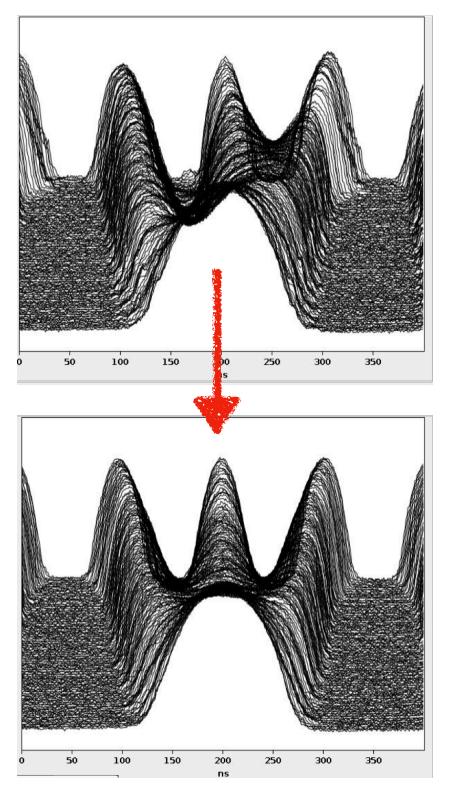








RL agent to correct **RF phase and voltage** to produce uniform RF splitting in PS for LHC beams

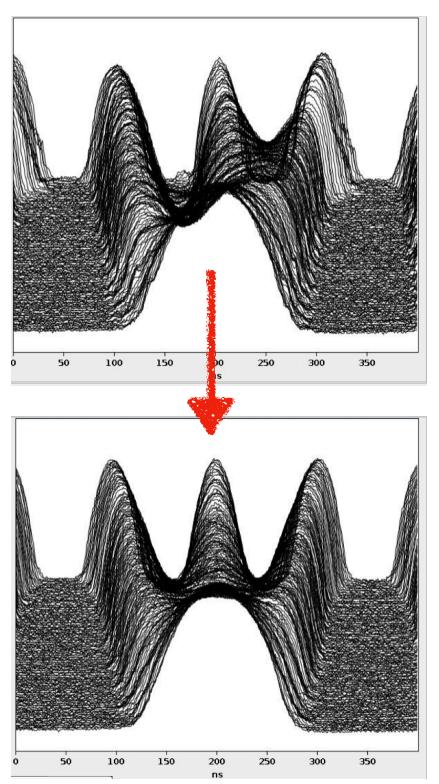


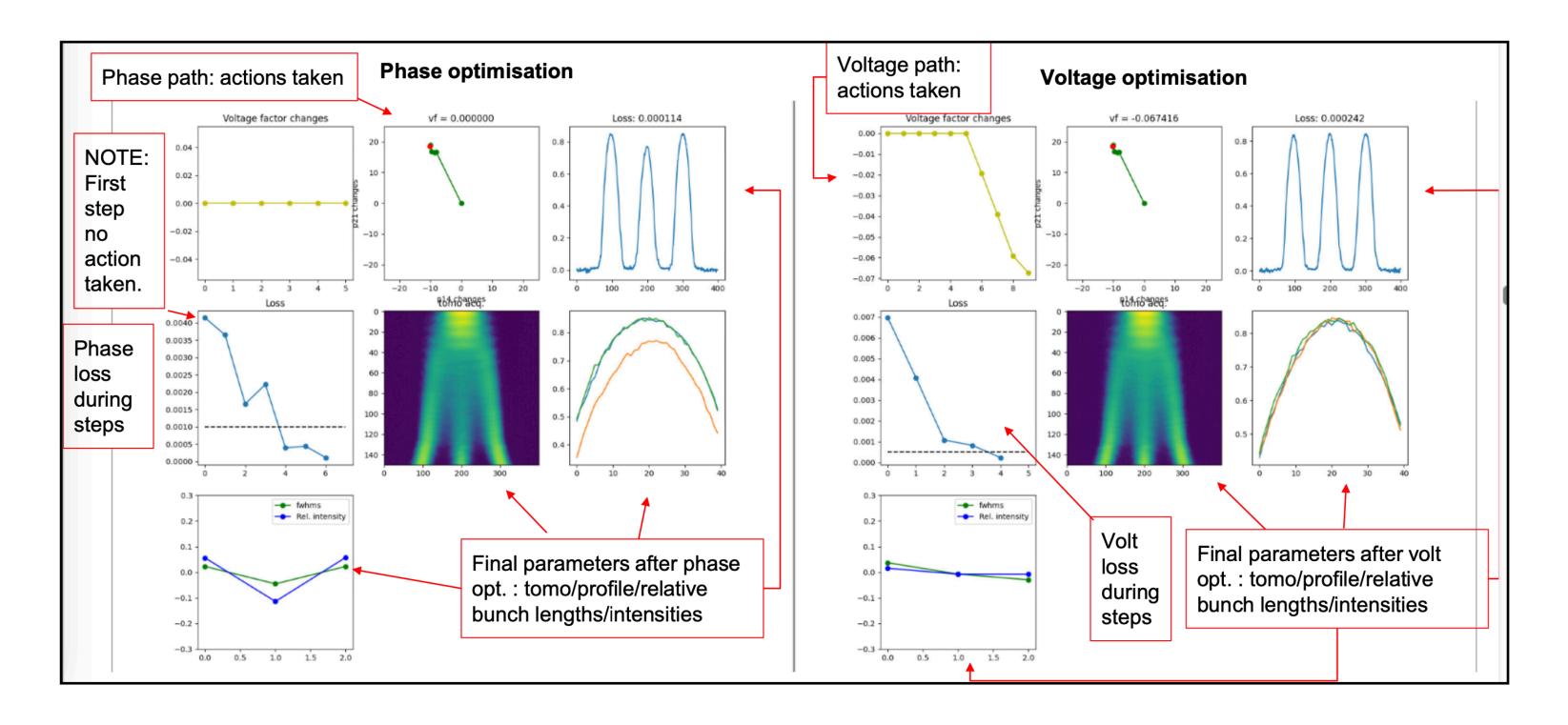




RL agent to correct **RF phase and voltage** to produce uniform RF splitting in PS for LHC beams

- Trained on simulation and successfully transferred to the controls room —> fully operational RL algorithm: Soft Actor-Critic (SAC); multi-agent algorithm using CNN to define initial set point Ongoing: from on-demand to continuous on UCAP









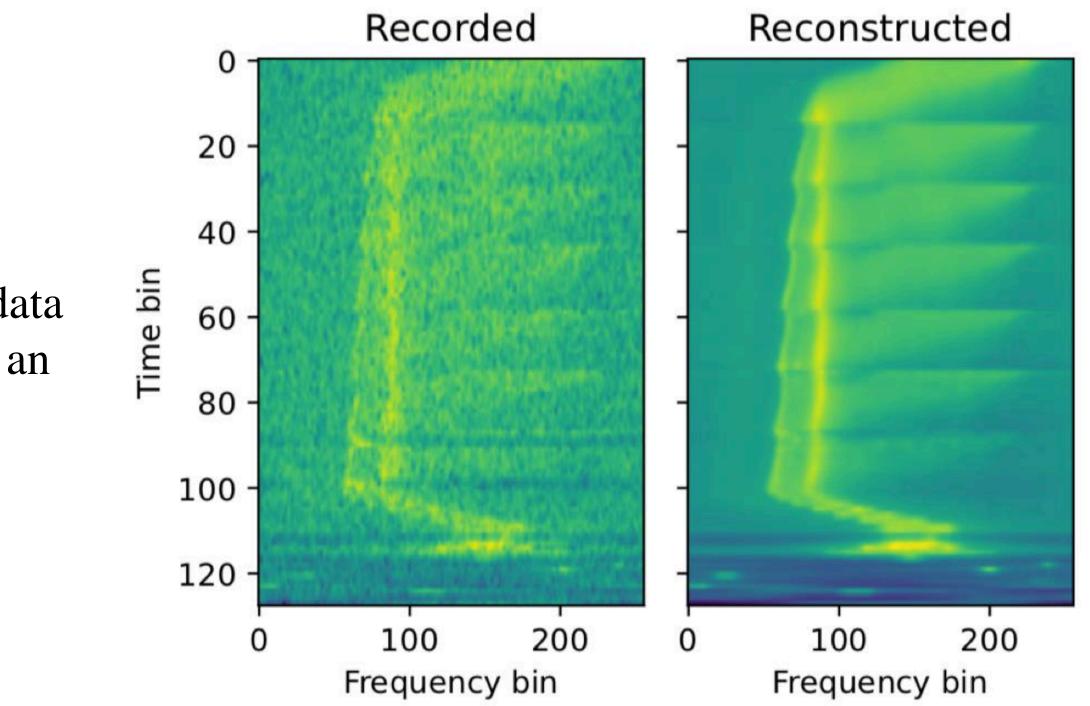
WIP: PhD ongoing to optimise injection efficiency into LEIR

- Control of ramping and debunching cavities in LINAC3
- Observation based on longitudinal Schottky spectrum
- Trained on data-driven dynamics model

Schottky spectrum from training data set (left) and its reconstruction by an variational auto-encoder(right)



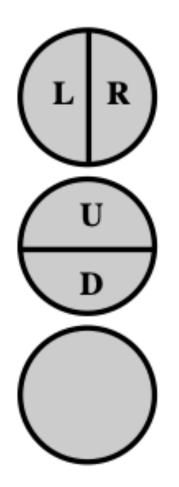




monitors (BSPs)

• RL state \vec{s} : $[(I_1 - I_2)_i]$ for each monitor; all intensities are normalised.

Our metric: symmetries per monitor: S



BSPH-BSMH Horizontal beam position

BSPV-BSMV Vertical beam position

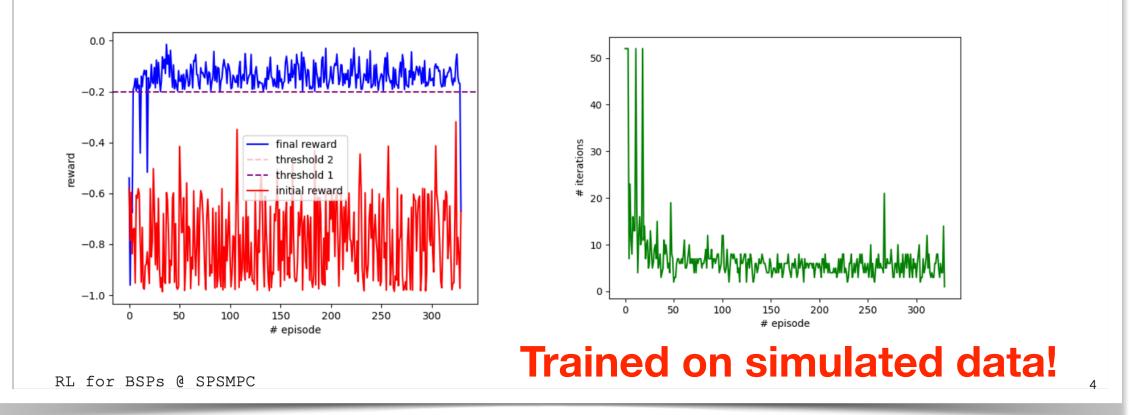
BSI Beam intensity

TT24: splitters to T4 - Training

• Test in horizontal plane: TD3; RMS reward

* Training time: 2000 iterations, 100 random steps

* Training emittance $2\mu m$, $100\mu rad$ random for initial trajectories



SX Workshop - Wr. Neustadt, A. Huschauer & V. Kain, 14-Feb-2024





Steering of DC beams in the CERN TT20 transfer line using split-foil secondary emission

$$= \sqrt{1 - \frac{|I_1 - I_2|}{I_1 + I_2}}; \text{ Goal: S > 0.8}$$

correctors = ["mssb.220450", "mbb.240428", bsps = ["bsph.240212", "bsm.241105", "bsph.241149"

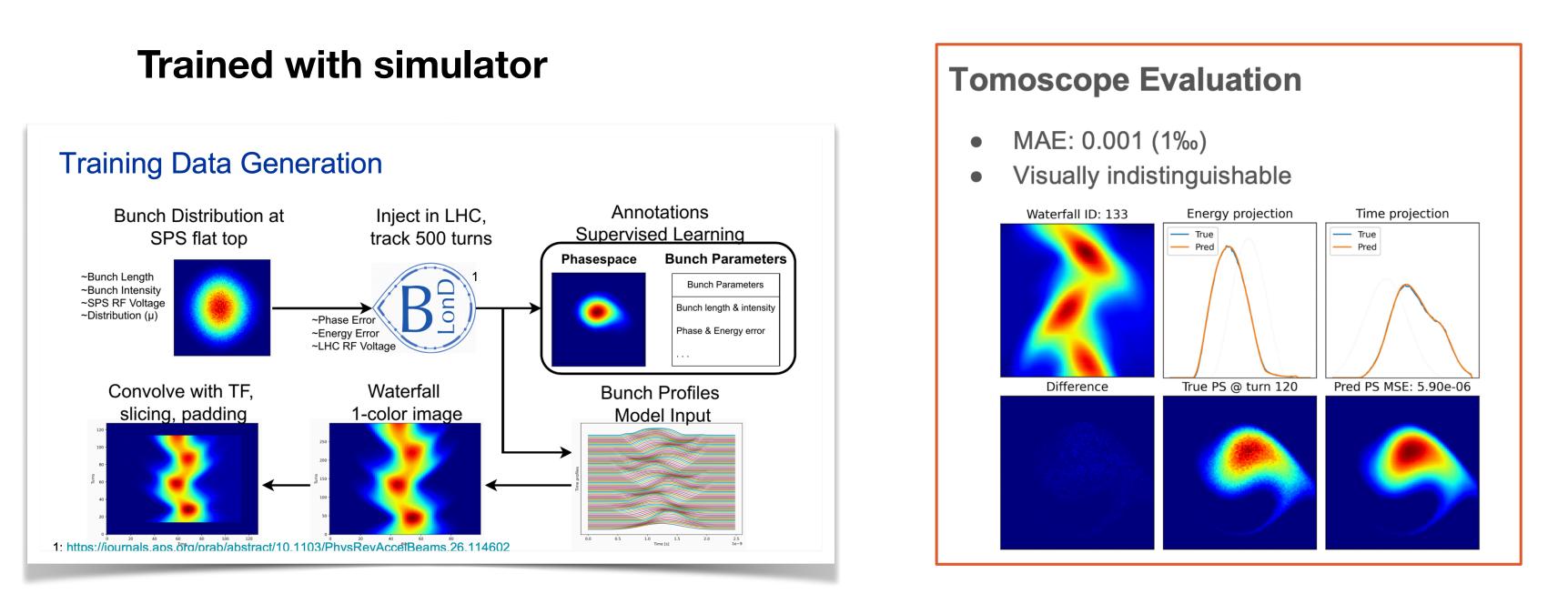
> Test of transfer foreseen for 2024 startup.

Enhancing diagnostics - ML tomoscope in the LHC

Speeding up tomographic bunch-by-bunch reconstruction in the LHC: using Auto-encoder ensemble: without AI bunch-by-bunch **not** possible.

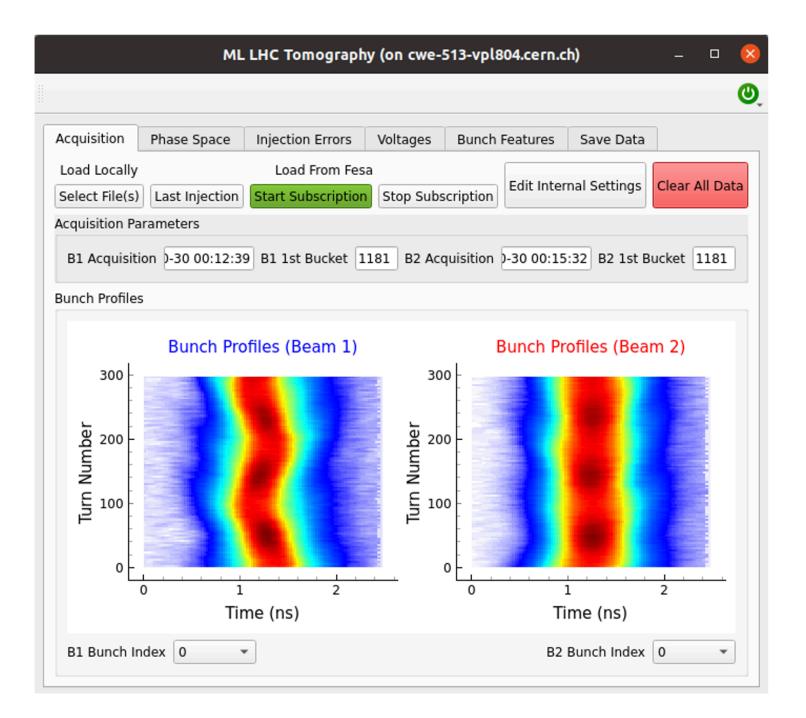
Fully operational on MLP

• Ready to be used in the LHC control room to measure injection errors





Courtesy K. Iliakis, T. Argyropoulos



Next steps: Efficient Particle Accelerators (EPA) project



10 work packages with interdependencies: work packages addressing technical topics and controls infrastructure evolution.

Each WP is high-level deliverable.

WP1 **Dynamic Beam Scheduling** Lead: F. Irannejad

WP4

Hysteresis Compensation Lead: C. Petrone

WP7 Automated Equipment Testing Lead: A. Calia, J.C. Garnier

WP0 Project Management Lead: V. Kain **Deputy: A. Huschauer**

WP2 Automated LHC Filling Lead: A. Huschauer, G. Trad

WP5 **Next Generation Sequencer** Lead: R. Gorbonosov

WP8 Automate Equipment Lead: F. Velotti, K. Papastergiou

SX Workshop - Wr. Neustadt, A. Huschauer & V. Kain, 14-Feb-2024



WP3

Automated Parameter Control & Optimisation Lead: M. Schenk

WP6

Efficient Settings Management Lead: M. Hostettler

WP9 **Data Processing Framework** Lead: M. Sobieszek

EPA goals

Focus is on automation of accelerator operation and equipment management • to increase efficiency, reproducibility, flexibility and hence performance





EPA goals

Focus is on automation of accelerator operation and equipment management

• to increase efficiency, reproducibility, flexibility and hence performance

scheduling beams

reduce impact on fixed target users and LHC turn-around time

and automatic containment of drifts

automate hardware commissioning

preparing the grounds for **preventive maintenance**



- WP1 **Dynamic Beam Scheduling: remove fixed supercycles** by algorithmically and dynamically
- WP2 Automated LHC Filling: automate and standardise LHC beam preparation and filling, and
- WP3 Automated Parameter Control and Optimisation: automation of parameter optimisation
- WP4 Hysteresis Compensation: deterministic field control with the aim of decoupling cycles
- WP7 Automated Equipment Testing: AccTesting for "all" equipment for injectors and LHC to
- WP8 Automate Equipment: automation of equipment setup, fault analysis, and recovery;

EPA timeline



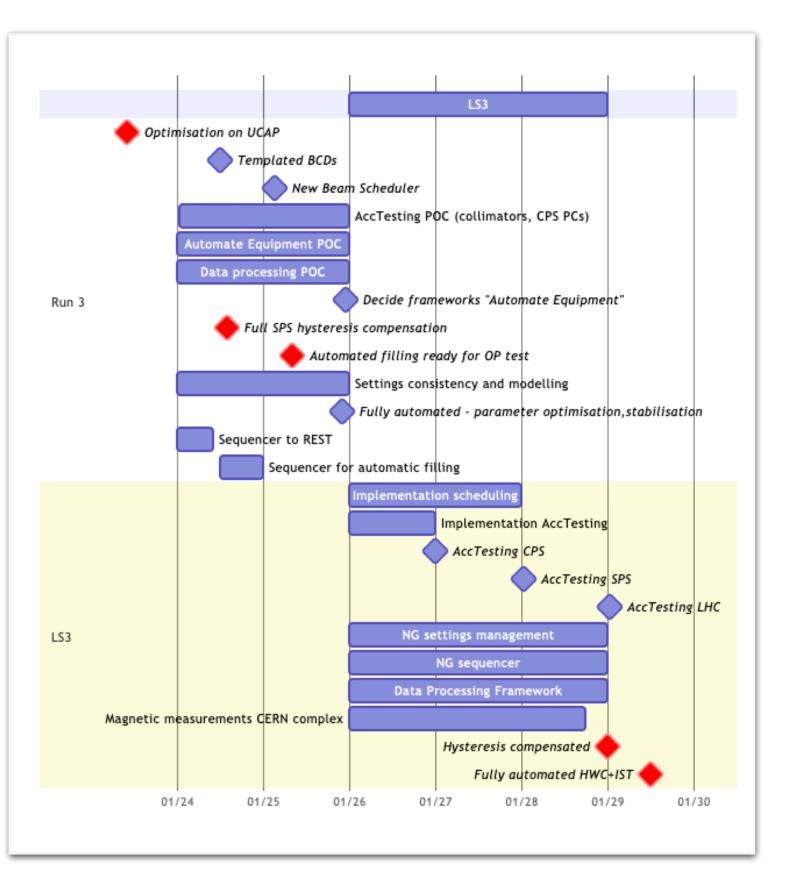
Time-bound project for 5 years

Improvements ready for Run 4 starting in 2029

Project with 2 phases:

- Prototyping and first operational tests in Run 3 (now!)
- Full implementation during LS3 and sequential commissioning in Run 4







Conclusions

Automation has long tradition at CERN





Conclusions

Automation has long tradition at CERN

The recent automation wave is radically different: data-driven!





Conclusions

Automation has long tradition at CERN

The recent automation wave is radically different: **data-driven**!

CERN has been investing into automation infrastructure and is working towards a paradigm shift with the **Efficient Particle Accelerators (EPA)** project

EPA was given 5 years to address various efficiency bottlenecks through

- Automation of accelerators and equipment
- Improved modelling
- Use of AI/ML at scale

