ISBN: 978-3-95450-209-7 ISSN: 2226-0358 doi:10.18429/JACoW-ICALEPCS2019-MOCPL02

MODERNIZATION OF EXPERIMENTAL DATA TAKING AT BESSY II*

R. Müller†, A. Balzer, P. Baumgärtel, O.-P. Sauer, G. Hartmann, J. Viefhaus
Helmholtz-Zentrum Berlin für Materialien und Energie GmbH, Berlin, Germany

Abstract

The modernization approach for the automation of experimental data taking at BESSY II will be based on the data model of devices. Control of new components and re-factoring and reassembly of legacy software should fit into a device based framework. This approach guides the integration of motors, encoders, detectors and auxiliary subsystems. In addition modern software stacks are enabled to provide automation tools for beam line and experimental flow control and data acquisition (DAQ).

Strategic goal is the mapping of real beam line components into modelling software to provide the corresponding digital twin. First tests applying machine learning (ML) methods within this context for tuning are promising.

MOTIVATION

The soft x-ray light source BESSY II recently celebrated two decades of user operation and one decade of forming the Helmholtz-Zentrum Berlin (HZB) [1] by merging the Hahn Meitner Institute (HMI) with the BESSY II facility. Thus HZB could offer two complementary experimental options for material science: the neutron reactor BER II and the synchrotron. BER II is shutdown by the end of the year 2019 and for the end of the next decade HZB intensifies plans of a successor soft x-ray source BESSY III.

Todays controls environment of the BESSY II experimental floor has been very much determined by the early days rapid installation of an extraordinary number of beam lines and instruments. To allow for most efficient use of the light source the organizational structure of BESSY II has been set up exceptionally flexible: many beam lines provided open ports ready to welcome complete user instruments; external institutes and cooperation research groups (CRG) could contribute by co-financing and co-operating beam lines and instruments. Ubiquitous switching mirror units (SMU) allowed to exploit x-ray source points in a time sharing manner. As a result BESSY II with only 14 usable straight sections could ramp up to some 35 beam lines and about 50 end station instruments in a very short time.

Two multi-purpose packages emerged and helped to keep the phase of heroic experiments limited despite the large variety of installations: a data acquisition (DAQ) hard and software set-up could cover data handling needs. Monochromator functions of beamline control rapidly developed into a large monolithic, multi-protocol, *generic* control program (called MONO) that enabled maintainability for the small group and the fast roll out of new beam lines. Even today this VME program also handles the communication to instruments, experimental flow control and insertion devices; sophisticated energy fly mode variants for dipoles and ID beam lines are implemented too.

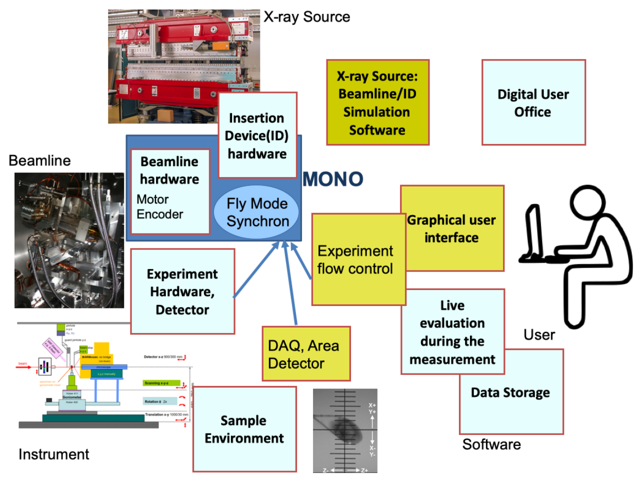

Figure 1: Sketch of the central role the multipurpose, generic software MONO (blue block) plays for the legacy experimental control infrastructure at BESSY

In consequence of a project oriented organizational structure of HZB new experiments are set up increasingly detector and method centric with instrument control, DAQ and experimental flow control planned and provided at a commissioning level with minimal resources. There is a free choice of the software tools needed (LabVIEW, IGOR, spec etc.), user friendliness and adaptability, failure tolerance and sustainability is not part of the project. In this low effort approach selection of the x-ray properties is again foreseen simply by remote control of the MONO software (Fig. 1). MONO has to integrate and adapt to new beam line components, e.g. hexapods, orchestrate with new insertion devices, handle obsolescence management of motors and encoders and provide handles to react on off-normal situation. As an VME program MONO is increasingly hard to maintain and develop on its generic level.

In general the control system structure in the experimental hall is by far too fragmented. Too many intertwined and interdependent connections need too many knowing hands to allow for easy and fast developments. At that point a clear strategic plan is inevitable to be able to over come the blocking constraints.

MODERNIZATION OPTIONS AND STRATEGY

Unlike other facilities BESSY II did not start with a holistic control system concept. Only for the accelerator a modern, consistent control system based on the EPICS framework was implemented. In the experimental hall no agree-

* Work funded by BMBF and Land Berlin
† roland.mueller@helmholtz-berlin.de

Content from this work may be used under the terms of the CC BY 3.0 licence (© 2019). Any distribution of this work must maintain attribution to the author(s), title of the work, publisher, and DOI.

17th Int. Conf. on Acc. and Large Exp. Physics Control Systems ICALEPCS2019, New York, NY, USA JACoW Publishing
ISBN: 978-3-95450-209-7 ISSN: 2226-0358 doi:10.18429/JACoW-ICALEPCS2019-MOCPL02

Content from this work may be used under the terms of the CC BY 3.0 licence (© 2019). Any distribution of this work must maintain attribution to the author(s), title of the work, publisher, and DOI.

ment on a common, overarching approach and planning existed, thus a very diverse collection of unique solutions emerged. Due to the large variety of interfaces and sub-control system functionalities and sizes today any modernization migration path is far from obvious.

A few larger building blocks exist and need to be analyzed. Most internal and external institutes run their LabVIEW, IGOR, Spec or custom software based systems for experimental flow control, data analysis and visualization. All rely on the flawless functioning of MONO. The experimental control department presently replaces the remaining EMP/OS 2/AT DAQ units with LISE/M /LabVIEW/PXI [2] based systems providing extended generic functionality, capable to cover a full experimental control environment. The sample environment department [3], originally serving the neutron experiments, works on a dedicated integration protocol, SeCop [4]. The MX experimental control is based on the common MX-Cube3 framework, standard at most facilities serving MX users. MX-Cube3 provides an autonomous, self contained, highly automated system, connected to beamline and accelerator control only by a few variables.

In that situation a general agreement on a common migration path along two core modifications appears to be crucial:

- Even if it sounds obvious: any unit, device, component of larger sub-systems integrated and "relevant" for user experiments or commissioning, needs to have an active and well maintained EPICS PV representation. It has to obey the HZB site wide, conflict free naming convention, in the beamline case a difficult endeavor (Fig. 2). This prerequisite is the key to unfold the full power provided by the rich EPICS framework: archiving, command logging, access control, state save/restore etc. *Connected* should become the standard equipment state. Thus at least the meta data describing the x-ray properties required for machine learning (ML, see section “Machine Learning Methods at the Beamlines”) and experimental conditions needed by FAIR [5] principles (see section “Novel Requirements”) will slowly become available.
- The beamline control software MONO needs refactoring: conditions of internal components, mirrors, grid, slits should be reported via EPICS PVs, properly named according to their physical meaning. The control part should still maintain consistency, protective actions and internal process tuning, while the component interplay should be easily adjustable by configuration data or scripts as commissioning or re-tuning findings develop (see overview figure in section “Beamline Optical Modelling” for the more general picture). In addition remaining compiled binaries should become independent from the VME/VxWorks platform.

COMPONENT INTEGRATION

Common EPICS connectivity of the installed devices within experimental set-ups to the command, analysis and surveillance software can help as a powerful architectural ordering principle.

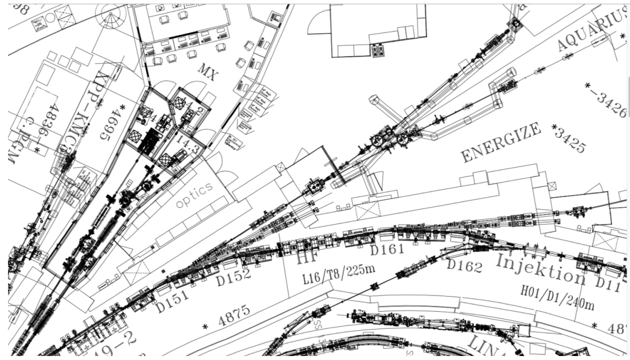

Figure 2: A small section of the experimental floor illustrates problems, that even arise for device naming. At BESSY sections of beamlines used in common split inside, outside or on both sides of the radiation protective wall. Dependent on the served instrument components in the common section change hardware representation, functionality and set point.

Simple devices are certainly trivial to integrate. Larger, complex components, like LabVIEW based solutions developed offline in the lab, or complete, encapsulated imported subsystems, that are based on an alien controls framework (TANGO, SECoP, etc.) are requested to provide appropriate EPICS connectivity to be fully integrable (Fig. 3). Similar arguments hold for slow controls of fast detector instruments, they feature custom internal DAQ data streams; and for PLC based subsystems; and for legacy configurations, where hardware control is done via detector specific (e.g. ArTOF, SES) or generic proprietary software (e.g. spec commands or scripts).

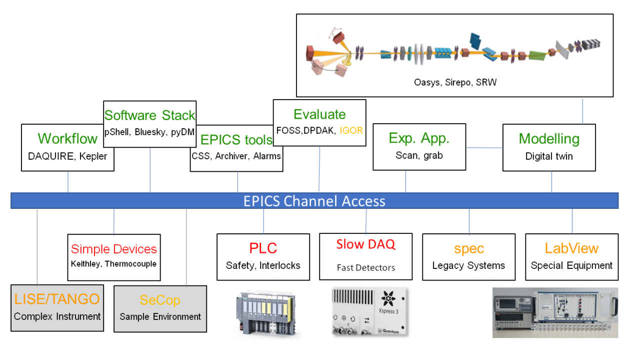

Figure 3: Plan for the refactoring, restructuring and reassembly of the experimental controls inventory. EPICS network representation has to be the unique integration layer for subsystems and software packages

For flat tools like archiver, alarm handling, operation logs collections of EPICS PVs will be adequate. Higher level tools, like sequencing, experimental flow control benefit from a device data model. Within the TANGO device server framework this comes for free. In the EPICS world the ophyd python package [6] provides device abstraction with

ISBN: 978-3-95450-209-7 ISSN: 2226-0358 doi:10.18429/JACoW-ICALEPCS2019-MOCPL02

desired functionality: a uniform specification for component hierarchies where devices have components, components may be other devices or signals, components can be categorized based on importance to operation. They can be operator level, access configuration attributes or control uncategorized properties. Simulations and test behavior without live devices is supported by the defined interface. Most common devices like motor, pseudo-motor based on motor record or image handling based on area detector are already implemented. Ophyd nicely fits to the babyIOC [7] base configuration.

EXPERIMENTAL FLOW CONTROL

As instruments became available at BESSY, DAQ software and experimental sequences have been largely based on students work, resulting in a mix of home grown solutions or proprietary software (IGOR, LabVIEW, spec). In view of limited maintainability and licence issues for off site data analysis, these implementations have to be transformed to a framework of tools that is based on Free Open Source Software (FOSS).

Early 2019 HZB hosted a workshop on *Automation in Beamline Control and Data Acquisition* [8] to be able to compare community supported tools with proven outreach and portability and to identify new trends and requirements of the experimental data taking process. For the EPICS based purposes of HZB the python stack of *ophyd* [6] device abstraction, *typhon* GUI generation based on ophyd, *Bluesky* [9] data collection, *pyDM* display manager together with packages like *transitions* for finite state machines appear to be most promising.

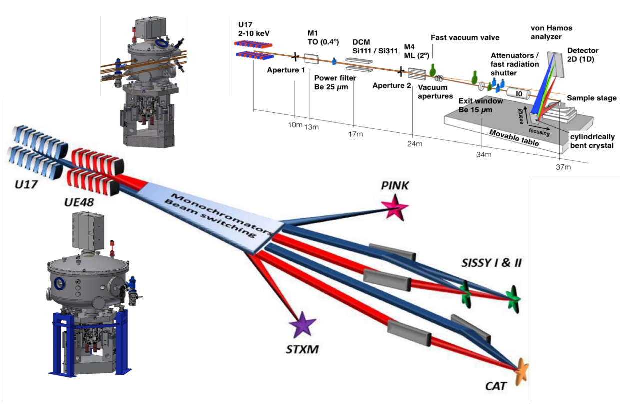

Figure 4: Two independent, canted undulators, two plane grating, one double crystal monochromator and multiple beamlines provide two color and polarisation combination x-ray capabilities to the Energy Materials In situ Lab (EMIL).

Additionally we will evaluate the multipurpose SCADA framework Sardana. This collaboration work of several European synchrotrons provides a similar device component based approach like Bluesky/Ophyd and features additional graphical tools for experiment control. In its current state it is mainly used in Synchrotron environments that use TANGO as low level control system although working experimental setups exist that integrate EPICS devices as well. Other solutions like the stand alone *pShell* [10] Java package can coexist with a caveat: device definitions within different solutions need to be carefully kept in sync.

Capabilities of these software packages allow to serve even the most complex beamlines, e.g. at EMIL (Fig. 4). There rapid, precise, coordinated switching is needed. Powerful sample robotics can feed the SISSY I + II endstations with fresh, well charactized samples as fast as every 4 minutes. Hard and soft X-rays have to be provided sequentially to probe the same sample spot at these and also at the CAT beamline. Two other beamlines, PINK and STXM beamlines are ready to utilize any left over beamtime. In total 4 competing end stations per source point have to be fed with variable wave length and polarization scan times. The intense, sequential usage of high power beams will lead to insufficient thermal equilibrium making active focus control mandatory for the given long beamlines (64m) and small focii (20μm).

HARDWARE TRIGGERED SCANNING

Precise coordination of high speed 2D detectors with mechatronic components allowing continous scanning on complex trajectories of the instruments with the photon pulses becomes increasingly mandatory. Ptychography resolution requirements ask for nano positioning capabilities.

Figure 5: PandA provides a common encoder processing platform based on Zynq 7030, supports multiple encoder standards (incremental, SSI, BISS...) and delivers synchronous triggering and data capture capabilities [11].

At other facilities simultaneous encoder processing, common synchronous triggering and data capturing for multi-technique scanning applications is provided by reprogrammable hardware units (PandA see Fig. 5 at Soleil/ Diamond/ MAX IV, Rythm at ALBA). At BESSY the advent of TimePix3 detectors and the short pulses of the BESSY-VSR (Variable pulse length Storage Ring) [12] upgrade are requesting these advanced, re-configurable synchronization components. With Malcom at Diamond and Flyscan at Soleil powerful role models for application frameworks exist.

BEAMLINE OPTICAL MODELLING

Commissioning, performance and process tuning as well as problem detection on beamlines benefit from the comparison of experimental results and their optical modelling. Both must be made accessible to the beamline scientist or the light

Content from this work may be used under the terms of the CC BY 3.0 licence (© 2019). Any distribution of this work must maintain attribution to the author(s), title of the work, publisher, and DOI.

17th Int. Conf. on Acc. and Large Exp. Physics Control Systems ICALEPCS2019, New York, NY, USA JACoW Publishing
ISBN: 978-3-95450-209-7 ISSN: 2226-0358 doi:10.18429/JACoW-ICALEPCS2019-MOCPL02

Content from this work may be used under the terms of the CC BY 3.0 licence (© 2019). Any distribution of this work must maintain attribution to the author(s), title of the work, publisher, and DOI.

source user in a simple and integrated way. Well established expert simulation tools, such as RAY-UI [13], PHASE and WAVE (in-house developments), or SHADOW, XRT, SRW, SPECTRA, OASYS, with their respective graphical user interfaces (GUI) must be provided reliably and automatically. Consistency of input parameters, such as positioning and type etc. of the optical elements and sources, will be guaranteed by a continuously maintained reference database (Fig. 6). It will be necessary to equip the beamlines with appropriate diagnostics to be able to improve model descriptions towards the real findings.

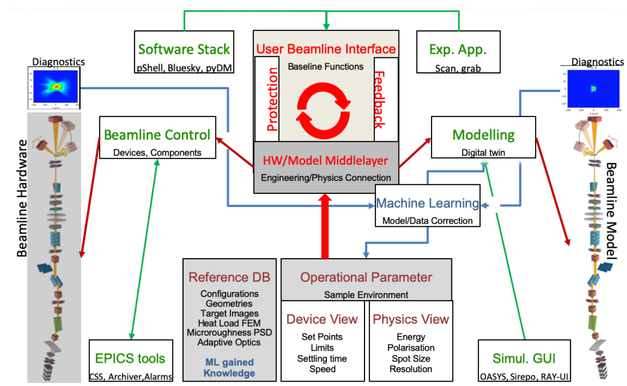

Figure 6: Sketch of the envisaged interplay of a high level user beamline control toolkit with configuration databases, device control, modelling and simulation software, diagnostics output and machine learning methods

BESSY aims at a high correspondence of model components and remote controlled hardware units. Strategic vision of this close mapping of beamline model and real implementation would be an ophyd device abstraction based middlelayer software (Fig. 6) that allows to connect tweaking actions w.r.t. model based performance tuning or fault detection with required operational parameter changes. In the accelerator world this is long time common practise.

Goal of the Australian Synchrotron (AS) project BRIGHT [14] is the development of control systems for vendor supplied beamlines for the model first and transfer it then to the real hardware as components become available. Their work break down structure could provide hints to a possible middlelayer structure.

Well adapted beamline models enable a drastic reduction of (re)commissioning and fine tuning times. Dry run capability of the model will contribute to the training of beamline scientists and users.

Today, the time-consuming process of refining the model by adapting the output data to the experimental results is no longer the only key to a good beamline understanding. The different methods of optical simulation can be supplemented and significantly accelerated by machine learning approaches.

MACHINE LEARNING METHODS AT THE BEAMLINES

Within the Helmholtz Association and the HZB there is a very general initiative to take advantage of machine learning (ML) methods, for process analysis and optimization and for the retrieval of scientific data from the measurements.

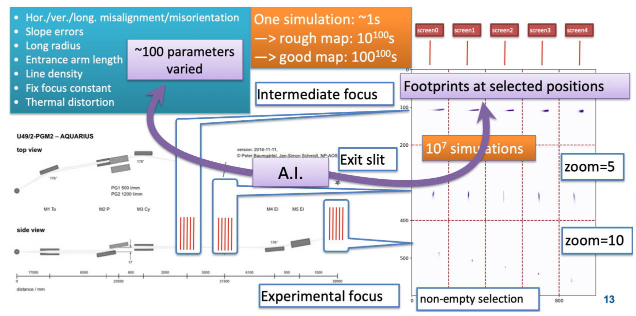

Figure 7: Process of model based generation of trained data for realistic beamline performance analysis.

At the light source BESSY II there is a close coordination of activities developing tool sets, such as training agents via Reinforcement Learning capable of optimising the experimental environment for the accelerator and the beamlines [15]. As pathfinder activities IT infrastructure prerequisites are studied, means of sharing methodology and correlating relevant operation parameters are developed. With the potential of the ML methods shortcomings of present model descriptions could be overcome and it might be possible to better describe source point properties and properly understand the x-ray preparation resulting from the actual settings (Fig. 7)

NOVEL REQUIREMENTS

Scientific data are a significant raw material of the 21st century. In support of the digital scientific methods the Helmholtz Association created the new topic and cross-program activity Data Management and Analysis (DMA). It aims at strong integration of modeling, simulation, experiment and analysis in a single, digital loop. It foresees development and application of state-of-the-art methods for forward and inverse problems (e.g. machine learning in imaging, simulations). The activity should provide applications, algorithms and techniques for in-situ, realtime analysis of data at high rates.

To add DMA values to the data taken by the experiments, capture and linked storage of meta data, describing the conditions of their experimental generation are required.

For light sources it would be a big step to have most beamline and instrument components remote controlled or constantly monitored, to be able to track relevant operational data within standard logging, archiving and snapshot data files. These data are prerequisite for DMA based process tuning of the property of x-rays usable for the experiment. Ultimately these data should also provide the means to reconstruct the experimental conditions for certain user studies.

ISBN: 978-3-95450-209-7 ISSN: 2226-0358
doi:10.18429/JACoW-ICALEPCS2019-MOCPL02

For the BESSY user community the consortium FAIR-mat [5] represents the interests of experimental, theoretical, and computational condensed-matter physics and materials science to make data Findable, Accessible, Interoperable, and Re-purposable (FAIR). The project NOMAD within FAIRmat has already made good progress [16] along the fourth paradigm of science (Fig. 8). Models for data policies, i.e. meta data systems for data taken at the large facilities are understood prerequisites for this new way of scientific exploitation.

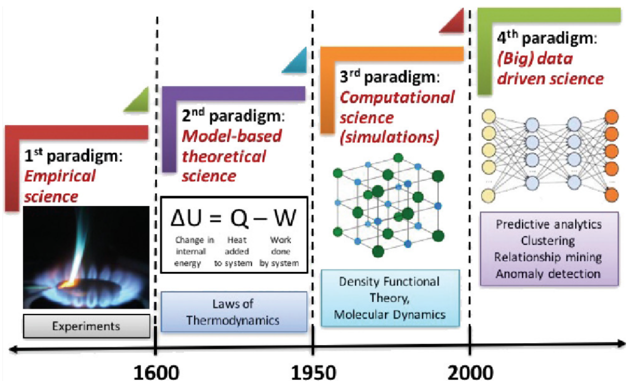

Figure 8: The four paradigms of science: empirical, theoretical, computational, and data-driven [17]. Scientists within the light source user community studying materials and condensed matter move to the 4th paradigm [16].

In Germany the National Research Data Infrastructure (NFDI), a proper, sustainable infrastructure for provision, inter linkage, maintenance, and options for reuse of research data in support of several consortia from a wide variety of scientific disciplines is under construction. This is well aligned to the European Open Science Cloud (EOSC) activities.

OUTLOOK

Modernization of data taking at BESSY is a long way to go with many stepping stones. It will be cumbersome and slow, since many world class instruments generate outstanding results and must not be affected by general upgrade activities. Nevertheless for a good future perspective a unified, sustainable data acquisition framework needs to be put in place. Handling of large data streams as well as hardware triggered scanning has to be enabled. Well adjusted automation tools should support efficient use of beamtime and provide data quality assurance via „near realtime analysis“. We have a clear vision, that should at least help to agree on the best way to proceed.

Facilities within the League of European Accelerator based Photon Sources (LEAPS) join efforts towards smart specialization: accelerators develop and push their specific expertise. Beamline and experiment controls tailor capabilities along interchanging, international user needs. User experience of controls and DAQ at BESSY will be compared with other LEAPS facilities. Today a complete „digital scientific workflow“ is at a very far horizon.

ACKNOWLEDGEMENTS

The authors thank all HZB staff members active in maintaining the complex controls infrastructure at BESSY II, including beamline and instrument scientists, accelerator controls, beamline optics as well as central IT networking and storage, for many fruitful and clarifying discussions.

REFERENCES

[1] 10 Years of Helmholtz-Zentrum Berlin, https://www.helmholtz-berlin.de/pubbin/news_seite?nid=20360;sprache=en;seitenid=74699

[2] O.-P. Sauer *et al.*, "LISE/M - A Modernised and Unified Modular Experiment Control System for HZB Beamlines", in *Proc. ICALEPCS'17*, Barcelona, Spain, Oct. 2017, pp. 1938–1940. doi:10.18429/JACoW-ICALEPCS2017-THPHA212

[3] HZB Sample Environment https://www.helmholtz-berlin.de/user/experimental-infrastructures/sample-environment/index_en.html

[4] SECoP — Sample Environment Communication Protocol, https://github.com/SampleEnvironment/SECoP

[5] FAIRmat, FAIR Data Infrastructure for Physics, Chemistry, Materials Science, and Astronomy e.V. https://fairdi.eu/fairmat2/consortium

[6] Ophyd https://nsls-ii.github.io/ophyd/

[7] babyIOC, Portable Control System in a Box https://oksanagit.github.io/babyIOC/

[8] Workshop on Automation in Beamline Control and Data Acquisition https://indico.helmholtz-berlin.de/conferenceDisplay.py?confId=11

[9] Bluesky Data Collection Framework http://nsls-ii.github.io/bluesky/

[10] PShell scripting environment for experiments developed at PSI, https://github.com/paulscherrerinstitute/pshell

[11] PandABox Project description https://www.ohwr.org/project/pandabox/wikis/home

[12] A. Jankowiak *et al.*, eds., "BESSY VSR – Technical Design Study", Helmholtz-Zentrum Berlin, Germany, June 2015. doi:10.5442/R0001

[13] RAY-UI: A user interface for RAY, https://www.hz-b.de/ray

[14] Standardisation of BRIGHT Beamline Control System, in *Proc. EPICS Collaboration Meeting*, Melbourne, 2018, https://events01.synchrotron.org.au/event/85

[15] L. Vera Ramirez *et al.*, "Adding Machine Learning to the Analysis and Optimization Toolsets at the Light Source BESSY II", presented at the ICALEPCS'19, New York, NY, USA, Oct. 2019, paper TUCPL01, this conference.

[16] NOMAD success stories, https://www.nomad-coe.eu/news-2/news-recent_success_stories

[17] Ankit Agrawal and Alok Choudhary, "Perspective: Materials informatics and big data: Realization of the "fourth paradigm" of science in materials science", *APL Materials*, vol. 4, p. 053208, 2016. doi:10.1063/1.4946894

Content from this work may be used under the terms of the CC BY 3.0 licence (© 2019). Any distribution of this work must maintain attribution to the author(s), title of the work, publisher, and DOI.