# Calibration and Calibration Data Processing Concepts at the European XFEL


S. Hauf, D. Boukhelef , P. Gessler, B. Heisen, M. Kuster, L. Maia, N. Raab, J. Sztuk-Dambietz, M. Turcato, K. Wrona, G. Weidenspointner and C. Youngman

European X-Ray Free Electron Laser Facility GmbH, Albert Einstein Ring 19, 22761 Hamburg, Germany



## Abstract

The European X-ray Free Electron Laser (Altarelli, 2006) is a high-intensity X-ray light source currently being constructed in Hamburg, Germany, that will provide spatially coherent X-rays in the energy range between 0.25 keV and 25 keV. The machine will deliver a unique time structure, consisting of up to 2700 pulses, with a 4.5 MHz repetition rate, 10 times per second at very high photon fluxes up to $10^{17}$ photons/s (Tschentscher, 2012). The LPD (Hart, 2012; Koch, 2013), DSSC (Porro, 2010, 2012; Lutz, 2010) and AGIPD (Graafsma, 2009) detectors are being developed to provide Mpixel imaging capabilities at the aforementioned repetition rates for a dynamic range spanning from single photon sensitivity to $10^4$ –$10^5$ photons per pixel. The detectors are optimized for specific energy ranges.

A direct consequence of the aforementioned detectors' characteristics is that they generate raw data volumes unprecedented in photon science, ranging up to 1Mpixel x 640 memory cells x 10 pulse/s x 16 bit, i.e. 12.8 Gbyte/s. On-detector vetoing may not necessarily lower these rates much - a memory cell freed by a vetoed pulse may be used by data from one of the remaining 2700 pulses a train consists of. The PC-layer may reduce this data amount by additional software triggering, but this is not guaranteed. Figure 1 gives an overview of the different data products at European XFEL, as well as their flows and involved user roles, under the assumption that processing takes place within XFEL's Karabo framework (Heisen, 2013).

In addition to the high data rates, the Mpixel detectors' on-sensor memory-cell and multi-gain-stage architectures necessary for the high dynamic range, pose unique challenges in detector-specific data corrections and calibration (Weidenspointner, 2012; Sztuk-Dambietz, 2013a). These challenges are addressed by providing a dedicated and thoroughly characterized set of test stands, which utilize continuous sources (Fe-55, X-ray tubes) as well as a pulsed setup: PulXar (Sztuk-Dambietz, 2013b), which is designed to produce X-ray pulses of 50-150 ns duration, within a 0.6 ms burst followed by a 99.4 ms gap. The radiation it produces thus closely matches the XFEL pulse structure. Additionally, simulation tools are being developed to assist in detector characterization (Bohlen and Joy, 2013).


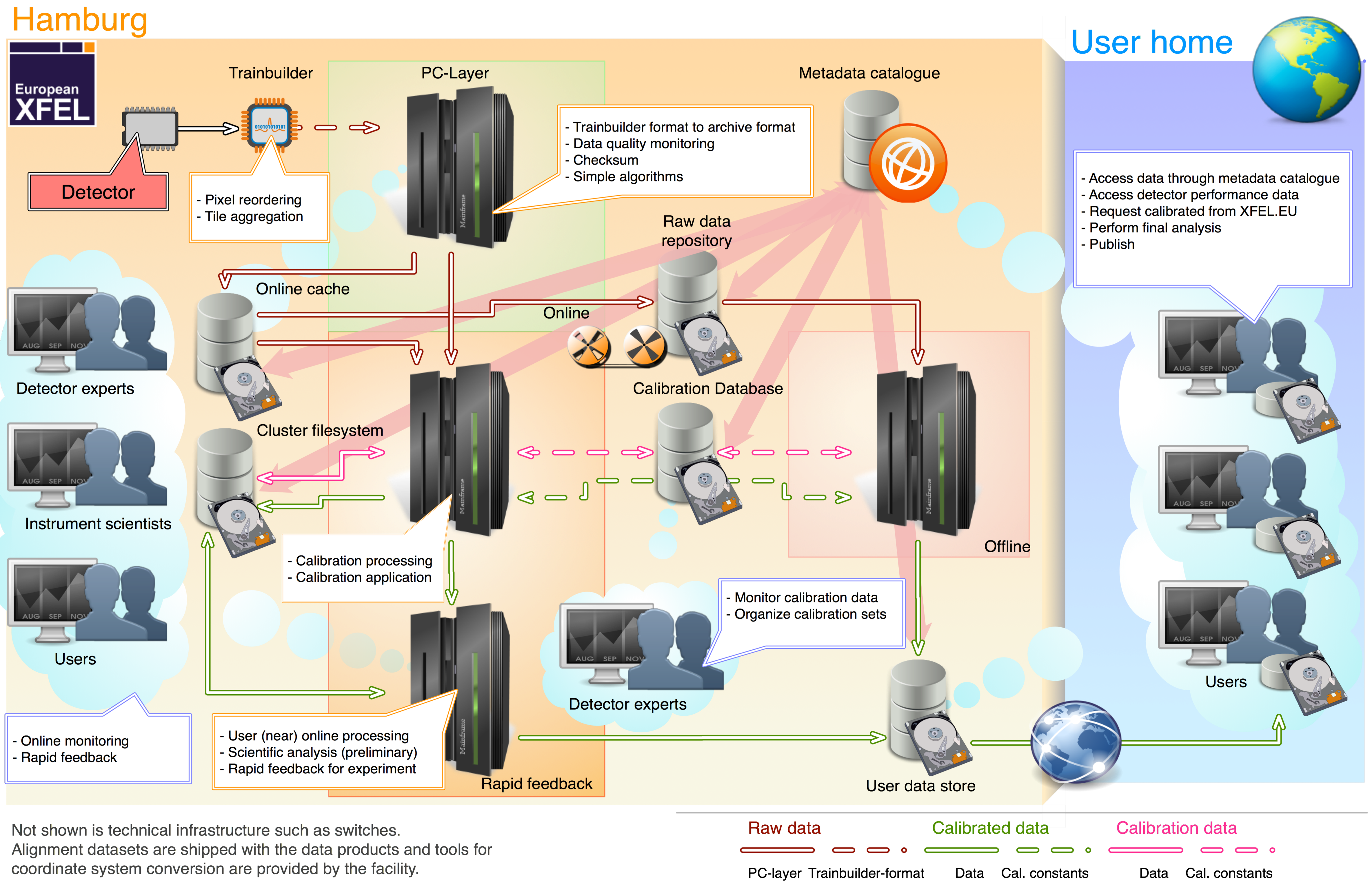


**Figure 1:** A general overview of data products, their flows, storage and related user roles at European XFEL, starting from data acquisition with a detector (top left corner). The different data products and formats are color-coded. Note how it is generally foreseen to provide users with calibrated data.

## Detector Analysis Suite

The large data volumes of the 1Mpixel detectors present challenges in two computational scenarios: (online) analysis and (online) correction and calibration. In the day-to-day work of the XFEL detector group the first scenario frequently occurs when characterizing detectors. Corrections may have only been partially applied, e.g. to investigate detector performance and its variation with operating conditions.

The latter scenario occurs in the context of standardized corrections in calibrated data production. The algorithms used here have usually been developed in the online scenario first.

We are developing a python library which is suited for both scenarios, usually running in iPython (notebook) (Pérèz, 2009) in the first or within Karabo devices in the second use case. Underlying, numpy and scipy (van der Walt, 2011) are used for calculations on the CPU, ipcluster for concurrent CPU processing and pyCuda (Klöckner, 2012) for GPU-accelarated computing. Documentation is realized in SPHINX. Table 1 gives and overview of the currently implemented functionality.

**Table 2**: Algorithms currently implemented in the analysis library

| Algorithm | CPU | GPU |
|---|---|---|
| Offset map calculation | X | |
| Noise map calculation | X | |
| Gain map calculation | X | |
| Streaming histograms | X | (X), partial |
| Statistics calculation | X | (X), partial |
| Offset correction | X | X |
| Gain correction | X | X |
| Common mode Correction | X | X |
| Pattern detection | X | X |

## Rel. Gain Characterization

The large number of calibration parameters (millions) poses a challenge to detector characterization algorithms in terms of throughput and stability. Gain characterization, e.g., is traditionally often performed by fitting a model of known photo peaks to calibration data. This approach works well for a small set of calibration parameters but can be problematic in terms of stability evaluation and throughput for a very large parameter space. We are evaluating an supplementary approach, based on a pdf-weighted mean estimator, which takes into account the probability $p_i$ of a given event $v_i$ belonging to peak $P_k$ or its neighbors $P_{k-1}$ or $P_{k+1}$ . The weight for each value $v_i$ is determined as

$$p'_{i,k}(v_i) = \frac{p_{i,k}(v_i)}{\sum_{k=-1}^{1} p_{i,k}(v_i)}$$

The mean peak location registered in a given pixel and memory cell is given by

$$\langle px_{x,y,c} \rangle = \frac{\sum_{i=0}^{N} p'_{i,k}(v_i) * v_i}{\sum_{i=0}^{N} p'_{i,k}(v_i)}$$

The stability of this algorithm with respect to detector resolution and peak estimate quality is shown in Figure 3. Note that this approach requires some a priori knowledge of expected peak locations determined e.g. through fitting of selected pixels.

## GPU Acceleration

GPUs are optimized for highly parallel processing, such as it occurs for per-pixel corrections of XFEL detector data. Using a Nvidia QUADRO K2200 (CUDA 6, compute capability 5) we are investigating whether real-time correction and calibration of streamed data is possible. The image sizes used for the test correspond to the physical tile sizes of the XFEL 1Mpixel detectors. Figure 2 shows example output from select currently implemented GPU algorithms, the correctness of which has been thoroughly tested.

Table 2 shows the currently achieved processing times per train. The available time window is 100 ms, which is currently exceeded by a factor of up to ~6 at a GPU resource utilization of 50–60%. Using more powerful TESLA hardware (e.g. a K20 with 4x more cores and memory bandwidth) and additional resource usage optimization, real-time corrections seem possible.

**Table 2**: GPU-based correction performance

| Detector (Pixels x mem. cells) | GPU (ms) | CPU (ms)** | Speed-Up |
|---|---|---|---|
| LPD (128x32x512) | 44.5 | 3473 | ~78x |
| AGIPD (128x512x352) | 521.3 | 15394 | ~30x |
| DSSC (256x128x800) | 651.1 | 19877 | ~30x |
| LPD Supermodule (256x256x512) | 810.2 | 35796 | ~44x |

** without pattern detection, 6(x2) core XEON W3680@ 3.33Ghz

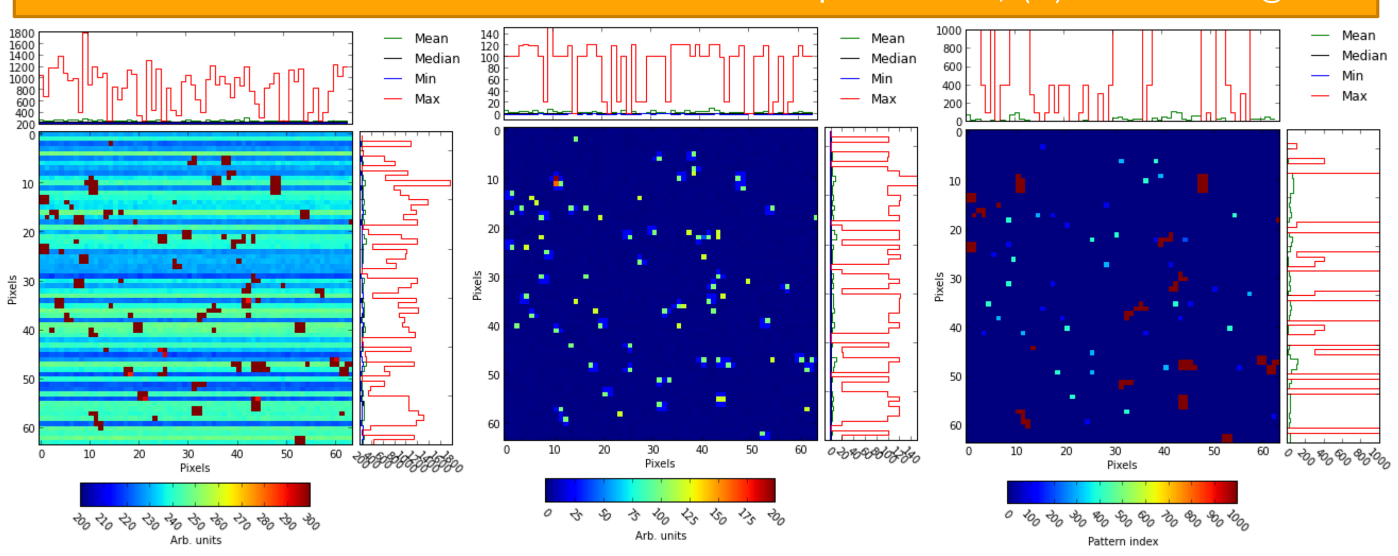


**Figure 2:** Examples of GPU-based corrections on test data. The left plot shows the uncorrected data with common mode effects highlighted; the center plot the same data after offset-, common mode- and gain correction have been performed on a per pixel, per-memory cell basis. The right plot shows the identified patterns (classified as singles (pattern index 100), doubles (200), triples (300) and quadruples (400)), as well as clusters (1000) with higher multiplicities.

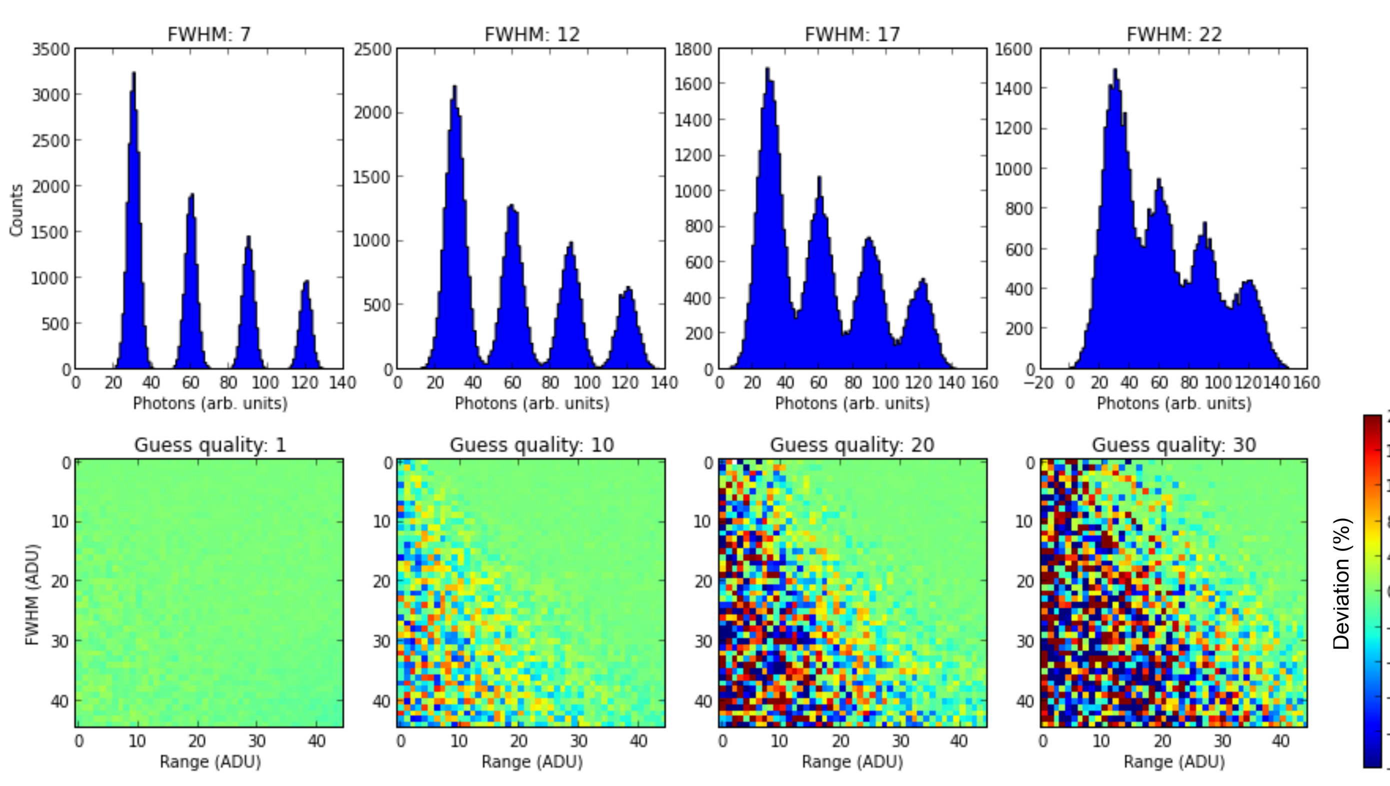


**Figure 3:** Top panel: examples of peaks with different FWHM which were used for testing the stability of the gain analysis algorithm. Bottom panel: gain determination accuracy in terms of deviation from the input peak location as a function of peak FWHM and range around each peak in which values are evaluated. The different guess qualities represent how well the location of a peak is known a priori (simulated by adding a random offset to the known peak location). Note how the algorithm is stable up to where guess quality and peak FWHM are similar (right-most plot).

## Calibration Database

Experts in European XFEL's detector group analyze the calibration data and the so obtained correction and calibration constants are injected into a calibration database. By assigning validity periods for each constant and providing calibration data updates as required, it is ensured that facility users will receive calibrated scientific datasets with standardized state-of-the-art corrections and calibration (Kuster et al., 2014).

The European XFEL calibration database runs on RDBMS and is accessible through RESTful API's and a Web Portal. It resolves calibration constant versioning in dependency on detector and experimental conditions (Figure 4). Currently, these conditions are aggregated in a management device, but will be drawn from a karabo-integrated run-management later. The calibration/ correction pipelines communicate with a Karabo device that in conjunction with a python library injects/retrieves the desired calibration (meta-) data from the database and the archive.

The initial implementation is completed and has successfully been tested on a unit-level and with test pipelines.

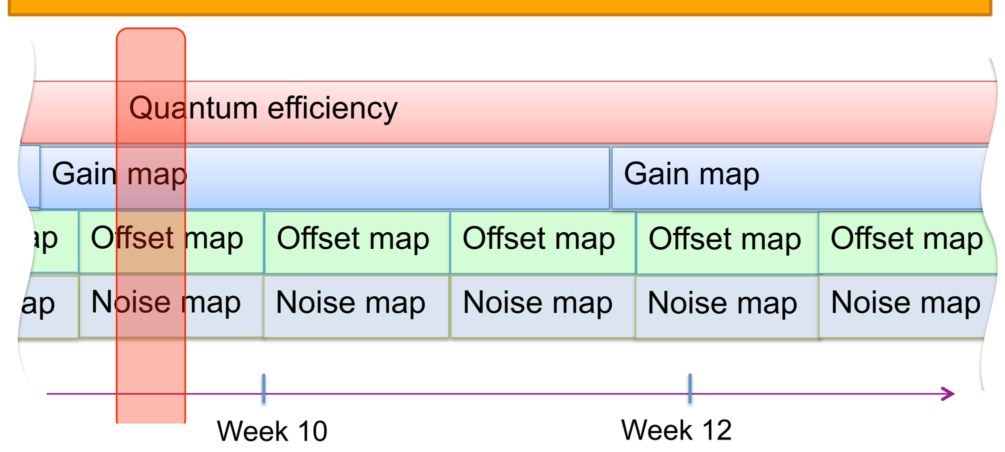


**Figure 4:** An example of how a requested detector calibration will map to different versions of different calibration constants.


### References

- Altarelli, M. et al., *European X-ray Free Electron Laser*, European XFEL GmbH, Technical Design Report (2006).
- Tschentscher, T. et al., *Photon beam properties at the European XFEL*, European XFEL GmbH Technical Report TR-2011-006 (2012).
- Graafsma, H., *Requirements for and development of 2 dimensional X-ray detectors for the European X-ray Free Electron Laser in Hamburg*, 2009 *JINST* 4 P12011.
- Porro, M. et al., *Development of a DEPFET sensor with signal compression: A large format X-ray imager with mega-frame readout capability for the European XFEL,* IEEE Trans. Nucl. Sci. Vol. 59, No 6 (2012) 3339.
- M. Porro et al., Expected performance of the DEPFET sensor with signal compression: *A large format X-ray imager with mega-frame readout capability for the European XFEL*, Nucl. Instrum. Methods, A624 (2010) 509.
- G. Lutz et al., *DEPFET sensor with intrinsic signal compression developed for the use at the XFEL free electron laser radiation source*, Nucl. Instrum. Methods A624, (2010) 528.
- Hart, M. et al., *Development of the LPD, a high dynamic range pixel detector for the European XFEL*, in Proceedings of the Nuclear Science Symposium and Medical Imaging Conference (NSS/MIC), 2012 IEEE, Oct. 27 – Nov. 3., 2012, Anaheim, CA, USA.
- Koch, A. et al., *Performance of an LPD prototype detector at MHz frame rates under Synchrotron and FEL radiation*, Journal of Instrumentation, Volume 8, Issue 11, C11001 (2013)
- Heisen, B.C et al., *Karabo: an integrated software framework combining control, data management and scientific computing tasks*, ICALEPS Conference Proceedings, San Francisco, CA, USA 2013
- Weidenspointner, G., et al., *Calibration of the non-linear system response of a prototype set-up of the DSSC detector for the European XFEL*, Nuclear Science Symposium and Medical Imaging Conference (NSS/MIC), 2012 IEEE. IEEE, 2012.
- Sztuk-Dambietz, J., et al., *Status of detector development for the European XFEL*, SPIE Optics+ Optoelectronics, International Society for Optics and Photonics, 2013.
- Sztuk-Dambietz, J., et al. *Laboratory Infrastructure for Detector Calibration Characterization at the European XFEL,* Nuclear Science Symposium and Medical Imaging Conference (NSS/MIC), 2013 IEEE. IEEE, 2013.
- Bohlen, M., Joy, A., et al. *Simulation Tools for Detector Performance and Calibration Sources at European XFEL,* Nuclear Science Symposium and Medical Imaging Conference (NSS/MIC), 2013 IEEE. IEEE, 2013.
- NASA, *Dictionary: SDP: Standard Data Processing – CIAO 4.5,* 10.6.2013. http://cxc.harvard.edu/ciao/dictionary/sdp.html (accessed 8.27.2013).
- XMM-Newton Science Operations Centre Team, *Users Guide to the XMM-Newton Science Analysis System,* ESA, 31.7.2013, http://xmm.esac.esa.int/external/ xmm_user_support/documentation/sas_usg/USG.pdf (accessed 8.27.2013).
- G. Lutz, et al., *An application of space technology to the terrestrial search for axions: the X-ray mirror telescope at CAST,* NIM A, 518.1 (2004): 201-206.
- L. Strüder, et al., *pnCCDs on XMM-Newton—42 months in orbit*, NIM A, 512.1 (2003): 386-40. In Science & Eng., 9, 3,.90-95, 2007
- Stéfan van der Walt, S. Chris Colbert and Gaël Varoquaux. The NumPy Array: A Structure for Efficient Numerical Computation, Computing in Science & Engineering, 13, 22-30 (2011)
- Fernando Pérez and Brian E. Granger. IPython: A System for Interactive Scientific Computing, Computing in Science & Engineering, 9, 21-29, 2007
- John D. Hunter. Matplotlib: A 2D Graphics Environment, Computing in Science & Engineering, 9, 90-95, 2007
- A.Klöckner, N. Pinto, Y. Lee, B. Catanzaro, P. Ivanov, A. Fasih, PyCUDA and PyOpenCL: A scripting-based approach to GPU run-time code generation, Parallel Computing, 38, 3,157-174, 2012
- Kuster, M., Boukhelef, D., Donato, M., Dambietz, J. S., Hauf, S., Maia, L., ... & Youngman, C. (2014). Detectors and Calibration Concept for the European XFEL. *Synchrotron Radiation News*, *27*(4), 35-38.





**e-mail: steffen.hauf@xfel.eu**