

Streamlining Processing and Utilization of EM Data - An Efficient Open-source Solution

Christoph Koch¹, Dieter Weber², Alexander Clausen², Andreas Mittelberger³, Sherjeel Shabih¹, Johannes Müller¹, Benedikt Haas¹, Alberto Eljarrat¹, Jonas Weinrich⁴ and Chris Meyer³

¹Humboldt-Universität zu Berlin, Berlin, Berlin, Germany, ²Forschungszentrum Jülich, Jülich, Nordrhein-Westfalen, Germany, ³Nion, Kirkland, Washington, United States, ⁴Ferdinand-Braun-Institut Leibniz Institut für Höchstfrequenztechnik, Berlin, Berlin, Germany

Modern transmission electron microscopes have become very complex instruments that are, in principle, capable of generating large amounts of very diverse information about a single sample. In addition to the microscopes themselves, current detector technology is evolving very rapidly, producing very large amounts of highly quantitative data in very short amounts of time. The high price of advanced microscope and detector hardware (often paid from taxpayer's money) makes the data acquired by these instruments highly valuable, calling for higher usage rates and broadening the use of the generated data. Although not being applied as much in materials science as in life science TEM, the technology for automating routine measurements is, in principle, available. Viewing these hardware developments in the context of the currently very rapidly evolving field of data-driven materials science (e.g. <https://materialsproject.org/> in the US or <https://nomad-coe.eu/> in Europe) there appears to be a discrepancy between the amount of available experimental data vs. the vast and quickly growing number of openly available computational materials data sets.

One prerequisite for data being useful to big-data driven materials science is accompanying metadata following some agreed-upon standard, that is well-suited to describe the actual data in such a way that it can be directly compared with or augmented by data acquired using other microscopes, very different techniques, or even theory. Another important aspect is the availability of software that implements a user-friendly and intuitive workflow for fast processing of large data sets to allow for on-the-fly decisions whether the acquired data is suitable for further processing and ingestion into a (local) data base from where it may ultimately be made available to data-driven materials science applications. This would largely increase the potential usability of data and produce a treasure that will continuously increase in value as current trends in the development of science suggest.

This talk will present recent developments of integrating the power of several open source software projects to create an intuitive user interface [1] that is capable of very efficiently processing arbitrarily large data sets [2] and ingesting experimental and derived data in an online accessible electronic lab notebook / searchable database [3] from where it can be made accessible to collaboration partners and other purposes. Initial examples with 4D-STEM data acquired in a modified SEM [4] and advanced processing of a spectrum image acquired in a Nion HERMES electron microscope [5] are shown in Figs. 1 and 2 [6]. Further examples will be presented during the talk.

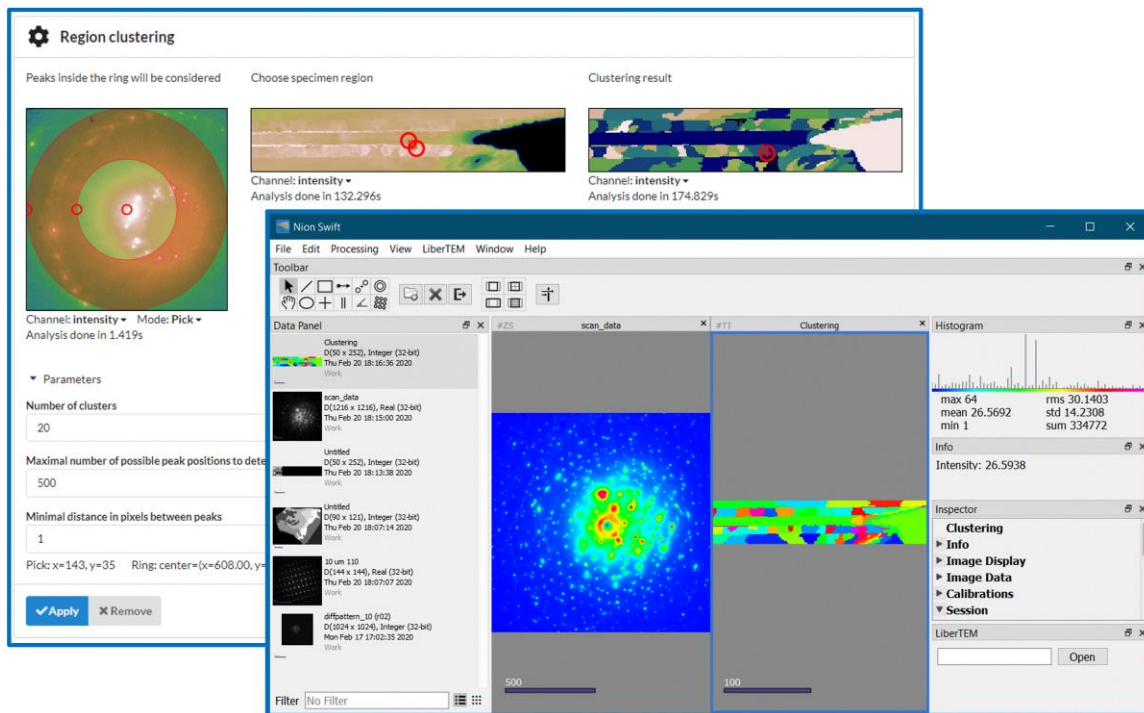


Figure 1. Two screenshots of a cluster-analysis performed once within the LiberTEM web-GUI and also within Nion Swift (performing the actual computation via the LiberTEM backend) by processing a 34.7 GB 4D-STEM Data set of AlN deposited on m-plane sapphire [6]. The data set was acquired in a GeminiSEM 500 (Carl Zeiss Micorsocpy GmbH) using a custom-built sub-stage that contains a fast fibre-optically coupled diffraction camera [4].

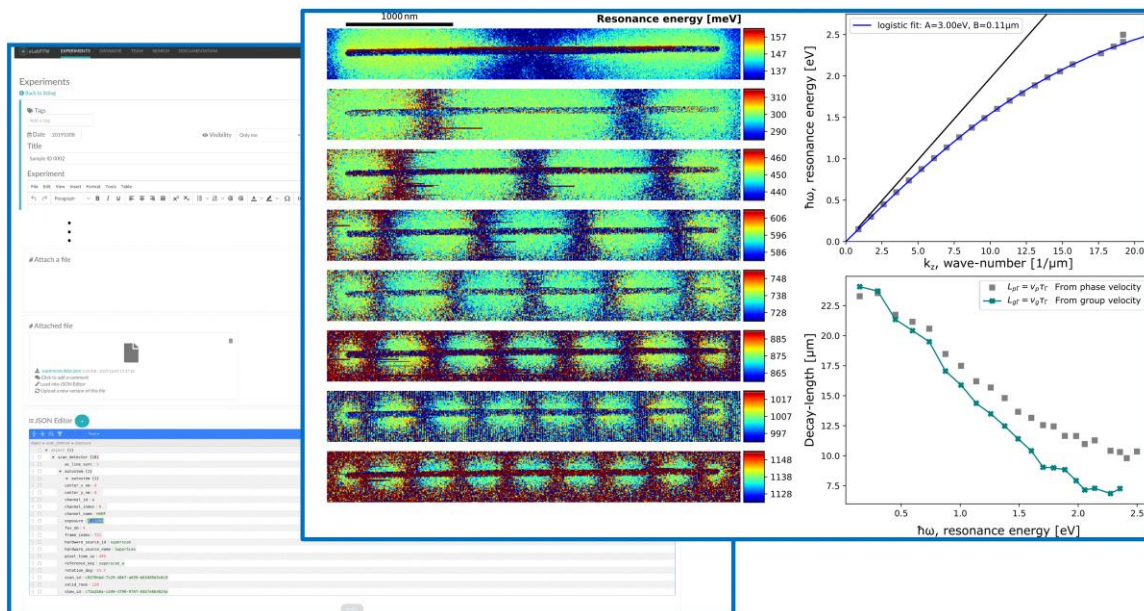


Figure 2. Background panel: Screenshot of the web-based electronic lab notebook elabFTW [3] that has been extended to display and edit metadata files in .json format. Front panel: example demonstrating the capability of the LiberTEM-library to process spectrum images of surface plasmon resonances in Ag nanorods [6] acquired using the Nion HERMES [5] at HU Berlin. The processing consisted in fitting a range of peaks to each spectrum in each of the pixels of the 512x64x2048 spectrum image. Left: column

of images corresponds to energy-resolved maps of energy loss to which Lorentzian peaks have been fitted in the range 0.1-1 eV. Within these 8 lowest energy modes we see some slight shifting of the energy position of these resonances. The right column shows 2 graphs, corresponding to the dispersion curve and decay lengths of these surface plasmon resonances. The decay-lengths were calculated using the measured peak broadening and estimated phase velocity $v_p = \omega/k$ and group velocities $v_g = d\omega/dk$ from the dispersion curve.

References

- [1] C. Meyer, N. Dellby, J.A. Hachtel, T. Lovejoy, A. Mittelberger, and O. Krivanek, *Microsc. Microanal.* 25 (Suppl 2) (2019) 122; open-source project: <https://nionswift.readthedocs.io/en/stable/>
- [2] The open-source LiberTEM project: <https://libertem.github.io/LiberTEM/>
- [3] eLabFTW, an open-source electronic lab notebook: <https://www.elabftw.net/>
- [4] J. Müller, B. Haas, W. Van den Broek, S. Fairman, and C.T. Koch, *Proceedings of MC2019*, IM5.004
- [5] O. Krivanek, N. Dellby, J.A. Hachtel, J.-C. Idrobo, M.T. Hotz, B. Plotkin-Swing, N.J. Bacon, A.L. Bleloch, G.J. Corbin, M.V. Hoffman, C.E. Meyer, T.C. Lovejoy, *Ultramicroscopy* 203 (2019) 60
- [6] We thank Dr. Duc Dinh (Nagoya University, Japan) for the sample of AlN on m-plane sapphire and Prof. O. Benson and M. Rothe (Humboldt University, Berlin) for the sample of the Ag nanowire.