# Analysis and Visualization of Collision Cascades Data

*U. Bhardwaj, H. Hemani, A. Majalee, M. Warrier*

*Bhabha Atomic Research Centre, Visakhapatnam, India*

As ever increasing computational power drives longer simulations generating more and more data, extracting knowledge from the data is becoming more relevant. Modern data exploratory techniques like supervised and unsupervised machine learning and reactive visualizations can provide effective ways to identify patterns and get new insights [1]. In the field of radiation damage, use of machine learning is restricted to using supervised learning for development of inter-atomic potential surfaces [2,3]. The potential of ML for exploratory data analysis using is now being explored [4,5,6]. **CascadesDB** lays the foundation for future efforts in this direction with a standard database of MD simulation data of collision cascades. **CSaransh** is an open source application that builds upon this and provides a whole suite of data exploration tools and modern interactive visualizations accessible directly from web.

**CSaransh** analyzes cascades database to explore patterns, correlations and comparisons of many new parameters such as sub-cascades, defect cluster shapes, cascade volume and surface area in addition to the essential properties like defects count, in-cascade clustering etc. It uses algorithms from the field of unsupervised machine learning, dimensionality reduction, computational geometry and other statistical techniques to find the new parameters while efficient methods have been developed for essential properties. The reactive GUI with modern web frameworks provides interactive visualizations and plots to explore the results qualitatively and quantitatively compare the different cascades or statistics calculated for a group of cascades across energies or elements.

The suite uses different tools for various tasks. C++14 is used to efficiently post-process big simulation outputs and has to be run only once for a given set of data. Python is used for the machine learning algorithms. HTML with React-js is used to develop the single page web application. JSON is used as the common data format between post-processors and GUI. The post-processing and web application both are modular such that new data-exploration parameters, plots and visualizations can be added seamlessly. *The initial results show effectiveness of new methods of data exploration to cascades and might open new avenues for better insights and possible refinement of higher scale models of radiation damage.*

**References:**

1. M. Ceriotti, The Journal of Chemical Physics, 150 (15) (2019) 150901.
2. J. Byggmästar, A. Hamedani, et al., arXiv preprint arXiv:1908.07330 (2019).
3. H. Wang, X. Guo, et al., Applied Physics Letters 114 (24) (2019) 244101.
4. U. Bhardwaj, et al., 4th MoD-PMI-2019 Workshop, June 2019, NIFS, Gifu, Japan.
5. U. Bhardwaj, AE Sand, M Warrier, arXiv preprint arXiv:1811.10923 (2019).
6. U. Bhardwaj et al.. Journal of Open Source Software, 4(41) (2019) 1461.