Editorial: Visualizing offline and live data with AI (VOLDA) workshop first edition Princeton 11-13th June 2024

D. Mazon[1]*, L. Abadie[2] and R. M. Churchill[3]

[1]CEA, IRFM, F-13108, Saint-Paul-les-Durance Cedex, France, [2]ITER Organization, Saint Paul Lez Durance, France, [3]Princeton Plasma Physics Laboratory, Princeton, NJ, United States

KEYWORDS

tokamak, artificial intelligence, fusion, machine learning, disruption

Editorial on the Research Topic
Visualizing offline and live data with AI (VOLDA) workshop first edition Princeton 11-13th June 2024

The first edition of the Visualizing Offline and Live Data with AI' (VOLDA) Workshop took place at the Princeton University Campus, Mader Hall from 11 to 13 June 2024. This annual workshop held for the first time aims at bringing together the fusion community to discuss the challenges brought by Artificial Intelligence (AI) and visualizing large datasets in fusion experiment and simulation.

Indeed, this subject becomes more and more important as the plasma durations in current fusion machines such as tokamaks increase, which range from a few tens of seconds to nearly 20 min. These long acquisition times generate a massive amount of information, notably images in the visible or infrared domains but also a very large number of physical measurements, to be exploited. In the perspective of ITER, the tokamak being built in Cadarache and for which the plasma duration will be up to 1 h, it is crucial to develop tools, which will allow rapid analysis of physical phenomena, anomalies, instabilities in order to protect the machine and avoid losses in performance of the fusion reaction.

The 3 days in hybrid mode event focused on feedback, lessons learned, and innovative techniques that developers and users experience with AI techniques and visualizing large datasets. The meeting was also a good opportunity for brainstorming about several open questions, decisions to be taken now for future application in fusion power plant.

Among the many discussed subjects, some main topics were addressed in details. We can cite.

- The smart indexing of dormant data and behind this how to visualize large data. It is clear that retrieving all points is impossible and thus down sampling techniques are crucial. The role of offline processing was also discussed and how to properly down sample a signal. The traditional way that consists on minimum/maximum/average may be not enough. The paper of Bhatia et al., "*Advanced Techniques for Fusion Data Visualisation*" presents innovative approaches and investigate how advanced visualisation might be adapted to work with fusion data, enhancing usability and integration.
- Also, how to handle a dynamic processing of large signals? Does this mean that the raw data processing in real–time is the correct solution? Some preliminary answers to these points are addressed in the papers of Castro and Vega, "*Smart decimation method applied to real-time monitoring*".
- How AI can help in the down sampling techniques and integrated data analysis? Concerning AI and its use for the operations, how to trust AI workflow and how to validate it? Churchill treat this problem in his paper "*AI foundation models for experimental fusion tasks*", with specific examples applied to fusion tokamaks experiments.
- The use of cloud systems in Fusion to store and to compute data. How the cloud is affecting data visualization? Shall everything run in the cloud? Which computing power is required? The papers of Amara et al., "*Accelerating discoveries at DIII-D with the integrated research infrastructure*" and Feibush et al. "*Visualization techniques for the gyrokinetic tokamak simulation code*" evocate this aspect and also describe actual techniques based on AI and ML to compute specific quantities in a dynamical manner to monitor and control the plasma.
- How to detect anomalies automatically is also treated from the mathematical point of view by Vega and Castro in his paper "*Automatic location of relevant time slices and patterns in both signals and video-movies: real-time and off-line visualization*" and by Boukela et al. in the paper "*Exploring NAS for Anomaly Detection in Superconducting Cavities of Particle Accelerators*".
- The place of open sources utilities: most visualization tools in fusion are proprietary. Defining a common github entry point could be a good starting point to collect all the tools. A next step would be to standardize the language; would Python be a right candidate?

All these points have been discussed in detail with some supportive presentation that were extremely appreciated. This meeting brought to light some particularly urgent questions to be solved. You will find in the associated papers many answers to these questions, of course not all of them have been fully answered and the next edition of the VOLDA meeting which is planned from 18 to 20 November 2025 at Madrid will complement the panorama.

Author contributions

Funding

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Generative AI statement

The author(s) declare that no Generative AI was used in the creation of this manuscript.

Any alternative text (alt text) provided alongside figures in this article has been generated by Frontiers with the support of artificial intelligence and reasonable efforts have been made to ensure accuracy, including review by the authors wherever possible. If you identify any issues, please contact us.

Publisher's note