Contribution ID: **104** Type: **Oral**

# On Intensity, Relativity, Shift & Compression –Video data archiving at W7-X

*Wednesday, 17 July 2024 11:40 (20 minutes)*

The Wendelstein 7-X fusion experiment is an optimized stellarator, equipped with more than 60 different diagnostic systems. Many of those diagnostics comprise optical and/or infrared spectrum sensor arrays of various technologies (cameras). In order to enable full compliance with accepted scientific methods, all data generated by these systems (during experiments) are stored in the W7-X ArchiveDB for 10+ years.

Prior to the operational phase OP2.1 conducted in 2022/23, all video data acquired by W7-X diagnostics has been stored in W7-X ArchiveDB as raw, uncompressed data. A growing number of cameras and camera based systems, rising frame rates and special resolutions induced challenges both in terms of network bandwidth and storage capacity. Increasing experiment run times to 30 minutes and above, combined with the ensuing necessity to ingest the video data into the archive in real time amplify these issues even further.

While other institutions chose to opt for storage of only selected data and doing so utilizing lossy of-the-shelf video compression schemes optimized around human perception, both approaches were deemed unacceptable by scientists at Wendelstein 7-X. After a detailed analysis of available open source and commercially available compression software, no pre-existing solution for all requirements could be found in the market. Consequently, researchers at W7-X teamed up with compression experts at Google Research to define a customized algorithm and data format. The core requirements were:

- to enable lossless retrieval of the originally acquired data,
- to achieve high throughput real time compression,
- to provide reasonable compression ratios, ideally below 50%,
- to be optimized for compressing the specific real-world data of the camera systems employed at W7-X,
- to well support the most common use cases in both data acquisition and retrieval at Wendelstein.

The result: Fusion Power Video.

This talk (and subsequent paper) describes the process of the development as well as the algorithms and data structures used by/with FPV and its integration into the W7-X archive infrastructure. In arguing the benefits and challenges of the solution, it is also a call for acquiring and processing generic camera data based on a first principles approach - to work with standardized, processing-optimized fixed point binary representations of relative intensity: left-aligned 16 (or 8) bit integers.

## Speaker's Affiliation

Max Planck Institute for Plasma Physics, Greifswald, Germany

## Member State or IGO

Germany

**Primary author:** FISCHER, Simon (Max Planck Institute for Plasma Physics, Greifswald, Germany)

**Co-authors:** VANDEVENNE, Lode (Google Research, Zurich, Switzerland); PERUSEK, Primoz (Cosylab, d. d., Control System Laboratory, Ljubljana, Slovenia); GRÜN, Martin (Max Planck Institute for Plasma Physics, Greifswald, Germany); HOLTZ, Andreas (Max Planck Institute for Plasma Physics, Greifswald, Germany); GRAHL, Michael (Max-Planck-Institut für Plasmaphysik); PRIBOŠEK, Jan (Cosylab, d. d., Control System Laboratory, Ljubljana, Slovenia); ALAKUIJALA, Jyrki (Google Research, Zurich, Switzerland); W7X-TEAM (Max Planck Institute for Plasma Physics, Greifswald, Germany)

**Presenter:** FISCHER, Simon (Max Planck Institute for Plasma Physics, Greifswald, Germany)

**Session Classification:** Data Storage and Retrieval, Distribution and Visulaization

**Track Classification:** Data Storage and Retrieval, Distribution and Visulaization