ALICE Data Preservation Status

- ALICE recognises the importance of Data Preservation, as an extension of the existing efforts to ensure availability of different generations of data and backward compatibility of analysis software for a large community of ALICE users

- We participate to the common discussions among LHC experiments and we embrace the general principles and guidelines that emerged from them

- We have not yet drafted a policy for DP, but have had several internal discussions on the subject
  - The MB is aware and endorsing this effort
  - We are trying to figure out from the technical and (in future) financial points of view what, when and how to implement the different levels of preservation, following the open access principle

Levels of DP in ALICE

Level 1: Publications, supporting documents and additional numerical data	Already in action, open-access publications, INSPIRE. We keep track of figures and additional data, to be propagated to INSPIRE and HEPData.
Level 2: High level, simplified data formats for simple analysis, theory comparison, education and outreach	We have tested and benchmarked vertexing analysis based on a very simplified event model. We are discussing making this format more general, to be used also by other ALICE analysis. The same format and analysis procedure can be then used for outreach.
Level 3: Reconstructed data and simulations, along with the analysis software environment , workflows and documentation to allow new analysis	We have a common analysis model and framework allowing to process in the same way all types of data formats suitable for analysis while the software is distributed via SVN. The same tools will be eventually made available for the public as currently only for the ALICE users – the discussions here are mostly centered on which data, how much resources and when (delays).
Level 4: Raw 'offline' data and the software and calibrations required to reconstruct them, including documentation	We preserve raw data (Tier0 + 1 copy on Tier1) for internal usage. The software used to process this data is open source and the documentation available. We do not foresee allowing large scale reprocessing of ALICE raw data by the general public due to the cost and considerable resources required.