Long term storage of number fields suffers from a hard problem: number fields do not have a canonical, unique representation, they are typically represented by an integral polynomial - which is non-canonical. Even worse, for databases, the size taken by such a presentation can differ by orders of magnitude.

Representation of invariants of the field, e.g. elements, thus in their presentation depend on the chosen polynomial of the field. This lack of unique-ness makes efficient databases difficult: retrieving data via keys and de-duplication have to depend on mathematically hard problems that do not have a fast (constant time) solution. This for example occurs trying to store representations of finite groups in characteristic 0 or Puiseux expansions of roots of bi-variate polynomials.

A similar problem is already visible in finite fields: while the field is uniquely specified by the size, the defining polynomial (the presentation) is not. There are several normalisations available in the literature (Conway polynomials for example), but typically they are extremely hard to get. Data depending on this, such as matrices defining codes or group representations, then have to be specified relative to the presentation.

In this context, any serialisation scheme need to support the notion of a parent of an object, providing the neccessary context in which the presentation of the elements can be understood.

In this talk, I will explain the problems and discuss some of the attempts we tried so far. In particular, in view of the OSCAR project and the data to be generated there, this is an important problem.