

# **Archiving Databases with SIARD**

**(Software-Independent Archiving of Relational Databases)**

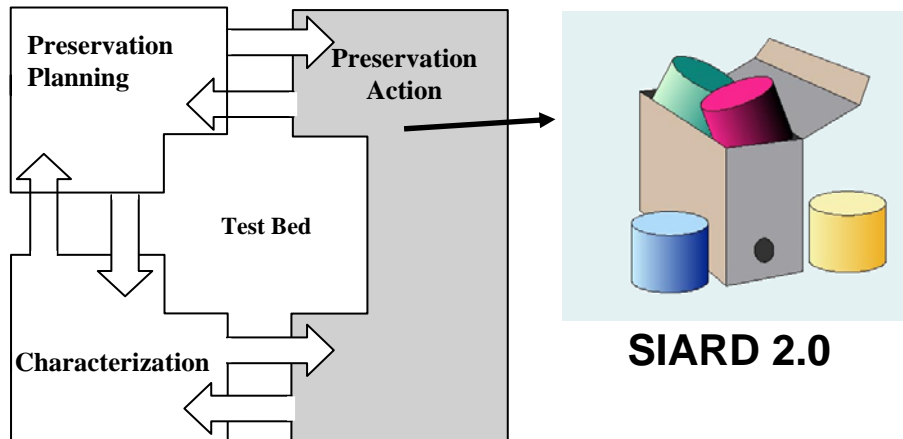
**16th International Congress on Archives, July 2008**

**Swiss Federal Archives**

Jean-Marc Comment



# SIARD within the PLANETS Framework



- SIARD command-line to be embedded in the PLANETS platform
- SIARD tools will be available in the future from the Swiss Federal Archives



# Challenges

- **M**ost administrative, scientific, business and official records are stored in databases. Today's challenge is making this information accessible for years to come and to ensure knowledge transfer as well as long-term administrative planning and stability.
- **L**ack of a standardized archiving format for database content has often rendered the preservation (i.e. archiving) task highly complex.
- **T**he Swiss Federal Archives (SFA) have taken up the challenge and developed a new long-term preservation format for a Software-Independent Archiving of Relational Database, short: **SIARD**.



# Solution: developing SIARD

- **The first SIARD prototype was initially employed by the SFA for archiving relational databases of government agencies. Since then SIARD has been considerably enhanced.**
- **The latest version (SIARD 2.0) is currently undergoing acceptance trials at the Swiss Federal Archives.**
- **The Swiss Federal Archives holds the copyrights for SIARD 2.0.**



# The SIARD Archive

- The SIARD format is based on internationally accepted standards.
- SIARD treats the most common type of relational databases. It enables structure (schemas, tables, etc.) and content of any given relational database to be stored in a simple XML coding. It can handle databases from varied provenances (e.g. MS-Access, Oracle and MS-SQL).
- A SIARD archive consists of a content file (content.xml) and a metadata file (metadata.xml) which includes metadata from all levels of the database. Both files are stored in a single uncompressed ZIP-container.
- SIARD is based on ISO standards (SQL:1999 and XML 1.0). The use of these standardized codes assures long-term preservation of databases archived in the SIARD format.

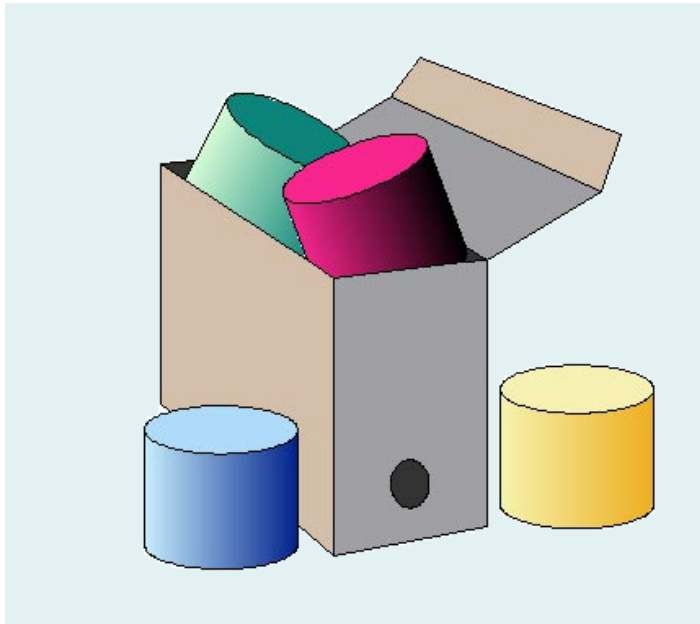


# User-friendly Archiving

- **Beyond the new format, SIARD also offers three central tools for archiving databases. The SIARD tools allow user-friendly metadata editing as well as access to the database content.**
- **SiardFromDb is a central command-line tool which extracts and converts a relational database into a SIARD archive.**
- **SiardEdit permits editing of metadata, matches it against those of a different archive, updates the metadata and views the primary data.**
- **SiardToDb uploads a SIARD archive into a database instance. It enables complex search queries to be conducted and navigation within a database.**
- **The JAVA based SIARD tools are platform-independent and can run on Windows, Linux or Apple systems.**



# SIARD Demonstration



**SIARD 2.0**



# Outlook

- The SFA's format was admitted as the new official archiving standard to the EU PLANETS project (Preservation and Long-term Access through Networked Services). This clears the way for international acceptance of SIARD. Archives, libraries, research institutes as well as private companies (e.g. IBM, Microsoft Research Limited, Tessella Support Services Plc) are involved in the PLANETS project. The prospects are favorable for the dissemination of SIARD throughout Europe.
- The latest format description is currently being prepared for publication. The definitive format report will be published in English as well as in Switzerland's three main official languages (German, French and Italian) The SFA intends to make the format description available on its website in September 2008.

