

— *Datenbanken für die Zukunft sichern*

Die SIARD Archivierungslösung für relationale Datenbanken



Schweizerische Eidgenossenschaft
Confédération suisse
Confederazione Svizzera
Confederaziun svizra

Eidgenössisches Departement des Innern EDI
Schweizerisches Bundesarchiv BAR

— Archivierung relationaler Datenbanken

02

Wichtige wissenschaftliche und administrative Daten werden häufig in relationalen Datenbanken¹ gespeichert. Dass diese Informationen auch in Zukunft zugänglich und nutzbar bleiben, ist von entscheidender Bedeutung. Dennoch war es bis vor kurzem nahezu unmöglich, relationale Datenbanken über lange Zeiträume zu archivieren, da es an entsprechenden Standards und Instrumenten mangelte.

Die Schweizer Lösung

Das Schweizerische Bundesarchiv liefert eine zweifache Lösung für das Problem: Zum einen entwickelte es ein offenes, publiziertes Archivierungsformat für relationale Datenbanken: das SIARD-Format (Software Independent Archiving of Relational Databases). Zum anderen stellt es ein Software-Paket namens SIARD Suite zur Umwandlung von Datenbanken in das archivfähige SIARD-Format zur Verfügung. SIARD-Format und SIARD Suite ermöglichen die Archivierung und langfristige Zugänglichkeit von Datenbankinhalten sowie den entsprechenden Metadaten² und Relationen.

Das SIARD-Konzept

Um sicher zu stellen, dass existierende Daten auch in Zukunft zugänglich bleiben, setzte das Schweizerische Bundesarchiv auf die Verwendung offener, publizierter und standardisierter Dateiformate. Das SIARD-Format und die SIARD Suite Software basieren auf international anerkannten Standards wie XML, Unicode und SQL:1999³. Damit garantiert die SIARD-Lösung eine wesentlich längere Lebensdauer für Datenbank-Inhalte als jedes nicht-standardisierte, proprietäre Format.

Das SIARD-Format

SIARD ist ein offenes Datenbank-Format, das speziell zur Archivierung entwickelt wurde. Die Spezifikation ist auf der Website des Bundesarchivs publiziert und kann kostenlos heruntergeladen werden. Das SIARD-Format ist unabhängig von der SIARD Suite-Software.

SIARD Suite

Die SIARD Suite Software konvertiert Datenbanken in leicht handhabbare XML-Dateien. Alle Primärdaten, Relationen und Metadaten bleiben erhalten. Der Benutzer kann die Primärdaten sichten sowie die Metadaten an die Anforderungen seiner eigenen Organisation anpassen und entsprechend ergänzen. SIARD Suite besteht aus drei Hauptanwendungen:

SiardFromDb ist ein Migrations-Werkzeug. Es konvertiert Datenbanken aus den drei am häufigsten eingesetzten Datenbanksystemen (Oracle, Microsoft SQL Server und Microsoft Access) in eine Datei im archivtauglichen SIARD-Format mit der Dateiendung «.siard». Diese neue Datei besteht aus mehreren XML-Dateien für die Primärdaten und aus einer XML-Datei für die Metadaten. Alle XML-Dateien werden in einem nicht komprimierten Zip-Ordner gespeichert (PKZip/Zip64).

SiardEdit ermöglicht den Benutzern die Metadaten zusätzlich zu dokumentieren und zu aktualisieren. SiardEdit erlaubt zudem eine Suche innerhalb der Metadaten. Die Benutzer können Primärdaten sichten und die Sichtanzeige nach verschiedenen Kriterien sortieren.

SiardToDb ermöglicht das Laden von SIARD-Dateien in alle unterstützten Datenbankformate: Oracle, Microsoft SQL Server oder Microsoft Access. So kann zum Beispiel eine Oracle Datenbank in eine SIARD-Datei konvertiert und dann als Microsoft SQL Server-Datenbank wieder geöffnet werden. Auf diese Art und Weise ist es möglich, erneut komplexe Suchabfragen in den archivierten Datenbankinhalten durchzuführen.

Nutzungsbedingungen

Das Schweizerische Bundesarchiv stellt die SIARD Suite Software gemäss dem gültigen Lizenzvertrag kostenlos zur Verfügung.

SIARD in Kürze

- Mit der SIARD-Lösung bietet das Schweizerische Bundesarchiv ein offenes, publiziertes Format für die Archivierung relationaler Datenbanken sowie ein entsprechendes Freeware-Paket namens «SIARD Suite», das einfach zu installieren ist.
- Das SIARD-Format basiert auf international anerkannten Standards wie XML, SQL:1999 und Unicode.
- SIARD Suite unterstützt die am weitesten verbreiteten Datenbankformate: Oracle, Microsoft SQL Server und Microsoft Access.
- Mit SIARD Suite lassen sich selbst sehr grosse Datenbanken archivieren.
- SIARD Suite ist Plattform- und Herstellerunabhängig. Es läuft in einer Java-Umgebung (1.5 oder höher) und wird unter anderem von Windows, Linux und Mac OS X-Betriebssystemen unterstützt.
- Die SIARD Suite-Benutzeroberfläche ist selbsterklärend und leicht verständlich. SIARD Suite ist auf Deutsch, Italienisch, Französisch und Englisch erhältlich.
- SIARD Suite wurde vom Schweizerischen Bundesarchiv entwickelt und vom europäischen Forschungsprojekt Planets⁴ als Standard in sein digitales Archivierungs-System übernommen. Das SIARD-Format wurde vom Schweizerischen Bundesarchiv zum offiziellen Datenbankenarchivierungsformat der Schweizerischen Bundesverwaltung erklärt.

Weitere Informationen, Bestellung der SIARD Suite Software und Anmeldung zu SIARD-Workshops unter www.bar.admin.ch

1) Eine relationale Datenbank organisiert Daten in Tabellen (Relationen), die untereinander verknüpft sind.

2) Metadaten beschreiben Inhalte und erklären die Bedeutung der Tabellen und Felder.

3) XML - www.w3.org/TR/xml/; Unicode - ISO 10646; SQL:1999 – ISO/IEC 9075:1999

4) www.planets-project.eu