

# RÉPUBLIQUE FRANÇAISE

Ministère de la culture et de la  
communication

## **Note d'information DGP/SIAF/2010/017 du 21 septembre 2010**

relative à l'étude du format SIARD pour l'archivage des bases de données relationnelles et au logiciel SIARDSuite mettant en œuvre ce format

Le directeur chargé des archives de France  
à

Mesdames les directrices des services à compétence nationale des Archives nationales

Mesdames et Messieurs les directeurs d'archives départementales  
sous couvert de Mesdames et Messieurs les présidents des conseils généraux

Mesdames et Messieurs les archivistes communaux  
sous couvert de Mesdames et Messieurs les maires

Mesdames et Messieurs les archivistes régionaux sous couvert de Mesdames et Messieurs les présidents des conseils régionaux

### **Annexes :**

- l'étude du format SIARD ;
- une note d'explications sur les concepts du modèle relationnel ;
- une liste de questions / réponses sur le logiciel SIARDSuite ;
- une description technique du jeu de test .

J'ai le plaisir de vous informer de la diffusion de l'étude sur le format SIARD et du logiciel SIARDSuite qui le met en œuvre.

Les Archives fédérales suisses ont fait développer dès 2004 un format spécifique pour l'archivage des bases de données relationnelles (SIARD<sup>1</sup>) ainsi qu'un logiciel (SiardSuite<sup>2</sup>) qui met en œuvre ce format dans un certain nombre d'environnements techniques.

---

<sup>1</sup> Software Independant Archiving of Relational Databases

<sup>2</sup> <http://www.bar.admin.ch/dienstleistungen/00823/00825/index.html?lang=fr>

Les archives de France ont souhaité étudier ce format et tester le programme associé afin d'évaluer leur possible utilisation dans le contexte administratif français et de déterminer si leur emploi pourrait améliorer les procédures d'archivage des bases de données relationnelles qui seront à prendre en charge pour archivage pérenne par les services d'archives. Ce sont les résultats de cette étude et de ces tests qui vous sont présentés dans l'étude en annexe.

L'étude distingue les observations faites sur le format de celles faites lors de l'utilisation du logiciel, tant sur des jeux de test que sur des bases réelles de production.

L'étude sur le format permet de mettre en valeur les caractéristiques pertinentes pour juger de sa capacité à représenter et à conserver de l'information.

Nous retiendrons de ces dernières que le format SIARD est un format dont la description est publique, qu'il est basé sur des formats standards et largement documentés et diffusés (XML et ZIP64).

Un des atouts de ce format est qu'il permet de récupérer d'une manière automatisée tant le contenu de la base que sa structure (liste des tables, description pour chaque table de la liste de ses champs, description pour chaque champs de son type). Le format permet aussi de décrire les relations entre les tables ainsi que des informations de gestion.

Ceci représente une amélioration notable des pratiques actuelles des services d'archives : archivage dit « à plat » des bases de données qui ne permet que la récupération des données seules, tables par tables, la documentation ( dictionnaires des données, nomenclatures, schémas des relations) n'étant prise en charge que d'une manière manuelle sans lien avec les données, ce qui en limite l'exploitation en particulier lors des communications.

Le logiciel SIARDSuite est un logiciel écrit en java et distribué gratuitement qui permet d'exporter une base de données dans un fichier au format SIARD ainsi que d'importer un fichier au format SIARD dans un SGBD<sup>3</sup>. La fonction d'importation a été développée pour permettre l'utilisation des SGBD comme outil de consultation.

Dans les atouts du logiciel SiardSuite, figure notamment la mise en œuvre du standard SQL-3<sup>4</sup>.

Toutefois, une limite de ce logiciel est qu'il ne supporte pour le moment que trois SGBD : Oracle, Microsoft Sql Server et Microsoft Access. Concernant ce dernier aspect, vous trouverez dans les annexes une description d'un test réalisé pour monter l'indépendance du logiciel SIARDSuite et du format SIARD<sup>5</sup>.

En synthèse, les caractéristiques du format SIARD lui confèrent une bonne capacité à représenter de l'information et à la conserver d'une manière pérenne. Son utilisation représente une amélioration de la pratique d'archivage actuelle des bases de données relationnelles. Les tests effectués sur des données réelles ont permis de confirmer ces atouts.

Je vous encourage vivement à prendre connaissance de cette étude, à expérimenter ce nouveau format<sup>6</sup> et à transmettre au bureau des traitements et de la conservation (sous-direction de la

---

<sup>3</sup> Système de Gestion de Bases de données

<sup>4</sup> Norme ISO/CEI 9075 : 1999 - Structured Query Language (SQL-3)

<sup>5</sup> Nous avons pour cela réalisé un script permettant d'archiver au format SIARD une base de données du SGBD mySQL (non supporté actuellement par SIARDSuite). Le fichier ainsi créé a pu ensuite être importé par SIARDSuite dans un des SGBD supporté.

<sup>6</sup>Le format SIARD et le programme d'archivage *SiardSuite* sont librement et gratuitement téléchargeables sur le site WEB des archives fédérales suisses à l'adresse suivante :

politique interministérielle et territoriale pour les archives traditionnelles et numériques), toutes les remarques que vous pourriez faire, qui pourront permettre d'améliorer le produit, sur la base du dialogue noué avec l'éditeur du format et du logiciel.

Le directeur

Hervé Lemoine