

SQL Server Internationalization

■ Overview

The SQL Server Internationalization workshop provides attendees with a broad understanding of internationalization processes, issues and pitfalls as well as the specific features and methods required to internationalize an SQL Server database. In particular, the issues related to full-text search on international data are presented.

The workshop shows how SQL Server deals with: character sets and Unicode (including surrogates, UTF encodings, normalization forms and transcoding), locales and locale models for client-server applications, resources and resource maintenance, sorting & searching, date & time processing, formatting of numbers and currency, text processing functions, etc. Schema internationalization (tables, indexes) is discussed along with its performance and scalability implications. SQL and Transact-SQL code internationalization is discussed for stored procedures, views, CHECK constraints, triggers and client-side SQL.

Attendees will leave with a clear understanding of how to correctly and efficiently internationalize their database schema and stored procedures.

■ Target Audience

This course is intended for DBAs, DB developers, software developers, software architects, software technical project managers and team leaders. It is highly recommended that attendees have a working knowledge of databases and SQL (and have taken the pre-requisite "All About Internationalization" workshop).

■ Benefits

This workshop provides DBAs and DB developers with a solid practical foundation on SQL Server internationalization. Time will be saved as this workshop presents a cohesive overview from many reliable sources along with tested, working examples.

■ Duration

The agenda described below is for a one-day session (there is also a 1.5-day version of this workshop with supervised hands-on exercises).

■ Pre-requisites

This workshop presumes that attendees have already taken the "All About Internationalization" workshop.

■ Agenda

1. SQL Server Internationalization

- SQL Server application development
- New globalization features in SQL Server 2008 vs. 2005 vs. 2000
- Useful books and articles

2. SQL Server and Locales

- Collations: the SQL Server locales
- Languages, character sets and collations
- Server, login and session languages: what they affect

3. SQL Server and Character Sets

- Character set architecture
 - N types are Unicode, "non N-types" are code pages
 - System values are "Unicode"
 - UCS-2 vs. UTF-16 (vs. GB18030 ?)
- Communications: ADO, OLE DB, DB-Library, transcoding
- Character set identification and enumeration
- Choosing character sets and a Unicode strategy
- Character set migration

4. Collations in SQL Server

- Collations: what they are and what they do
 - Sorting
 - Matching
 - CharSet for non-Unicode columns
- Case insensitive, accent insensitive, kana-insensitive and width-insensitive collations
- Windows collations and SQL collations
- Collations at the server, database, column and expression levels
- Collation Precedence Rules
- Indexes and multiple language indexes for a single column

5. Internationalizing the Schema

- The cultural dimensions & field cultural categories
- Field expansion, splitting, merging
- Normalization vs. culture
- Schema modifications for translation maintenance

6. Internationalizing the SQL Code

- Unicode data types and type conversions
- Text processing functions, stored procedures
- Date & time formatting
- Number and currency formatting

7. Full-text Search vs. Multilingual Data

- Languages and indexing
- Languages and querying
- Word breakers and performance

■ Handouts

Each attendee will receive a 220+ page booklet, one slide per page, with ample room for notes, complete with table of contents and glossary. The booklet is designed as a practical easy-to-use reference “book” for regular use during an internationalization project.

■ About our Instructor – Pierre Cadieux

Pierre Cadieux is a veteran with over 30 years' experience in internationalization of software, Web sites and embedded systems. He has taught internationalization at the Université de Montréal. Pierre has been technology editor for the LISA newsletter, VP Technology at ALIS and director of technology at Bowne Global Solutions.

At ALIS, Pierre pioneered the transparent handling of Arabic and Hebrew languages and created the core bi-directional technology licensed by Microsoft.

As Director of Localization Technology at Bowne Global Solutions, he carried out research and analysis on multilingual Web sites and published the first generic model of Globalization Management Systems.

Additionally, Pierre holds a B. Sc. and M. Sc. in Computer Science.