



VAMDC Query Language Specification

Document Information

Editors: L. Nenadovic, G. Rixon
Authors: G. Rixon
Contributors: Thomas Marquart
Type of document: standards documentation
Status: draft
Distribution: public
Work package: WP6
Version: 11.05
Date: 08/06/2011
Document code:
Document URL: http://www.vamdc.org/documents/querylanguage_v11.05.pdf

Abstract: VAMDC SQL Subset 1 (VSS1) is a small subset of SQL92 specified for the VAMDC data-services. This document defines the VSS1 language and its processing.

Version History

Version	Date	Modified By	Description of Change
V0.1	11/01/2011	G. Rixon	first draft
V11.5	27/05/2011	L. Nenadovic	editing for first release

Disclaimer

The information in this document is subject to change without notice. Company or product names mentioned in this document may be trademarks or registered trademarks of their respective companies.

All rights reserved

The document is proprietary of the VAMDC consortium members. No copying or distributing, in any form or by any means, is allowed without the prior written agreement of the owner of the property rights.

This document reflects only the authors' view. The European Community is not liable for any use that may be made of the information contained herein.

Acknowledgements

VAMDC is funded under the "Combination of Collaborative Projects and Coordination and Support Actions" Funding Scheme of The Seventh Framework Program. Call topic: INFRA-2008-1.2.2 Scientific Data Infrastructure. Grant Agreement number: 239108.

CONTENTS

1	VSS1	1
1.1	Introduction	1
1.2	Narrative definition (normative)	1
1.3	SQL92 syntax (informative)	2
2	References	4

VSS1

1.1 Introduction

VAMDC SQL Subset 1 (VSS1) is a query language designed for the VAMDC-TAP web-services. A query in VSS1 defines an extract of an archive that a data service can return in an VAMDC-XSAMS document or in a tabular format.

1.1.1 Subset of SQL92

VSS1 is based syntactically on ISO SQL92 but discards almost all the features of that language. The only features remaining are those relevant to selecting data to build an VAMDC-XSAMS document; VSS1 contains no feature for modifying a database. VSS1 assumes VAMDC's standard view of the database as a single table. This latter point means that there is only one visible table (which therefore does not need to be named in a query) and the column names are terms taken from the VAMDC dictionary. This is an example of a VSS1 query:

```
SELECT ALL WHERE AtomNuclearCharge = 25 AND AtomIonCharge < 2
```

In the following definitions, VSS1 queries are assumed to be submitted by a client application to a web-service. That service contains a query processor that converts the VSS1 query into a SQL query or set of queries, suitable for the relational database managed by the service. A request to the service contains the VSS1 text of the query and other parameters such as the desired formats of the returned results.

1.1.2 Interoperability and extensions

VSS1 is intended to be interoperable across all VAMDC databases. The same query can be accepted at all database, even though different data are raised from each.

Making VSS1 fully interoperable and capable of easy implementation would exclude all the higher functions of SQL (mainly features of the WHERE clause, such as sub-queries). Where these features can easily be added to a particular query-processor it would be a beneficial to have them. Therefore, certain parts of SQL92 are noted as extensions to VSS1. A query containing extensions is not strictly valid VSS1 and therefore not interoperable across all database; it may work on particular database. Users and applications should normally use only valid VSS1, but may use extensions on queries written specially for individual databases.

1.2 Narrative definition (normative)

To be a valid in VSS1, a query must satisfy the syntactic rules of the SQL92 language.

VSS1 queries never alter the database to which they are applied.

- VSS1 query must be a SELECT statement.
- VSS1 queries must not contain the SQL92 keywords ALTER, CREATE, DELETE, DROP, INSERT, REPLACE or UPDATE.

- VSS1 queries must not contain the SELECT ... INTO construction of SQL92.

VSS1 queries must not contain the JOIN keyword.

Column names used as operands in a VSS1 query must be terms taken from the VAMDC Restrictables dictionary. These column names must not be qualified by the name of a schema or database.

All the terms in the dictionary are valid as column names on all databases with a VSS1 processor. The query processor must implement the translation of the dictionary terms to names of real columns in the underlying database.

VSS1 processors may accept only a sub-set of the dictionary keywords, corresponding to the content of the underlying database. This sub-set naturally varies between databases and the set of restrictables for a given database is normally made available to the clients of the database. Where a query includes restrictables not supported by a given VSS1 processor, the processor must reject the query; it must not process the query while ignoring the unsupported constraints.

When the results of a query are to be returned in VAMDC-XSAMS format, VSS1 queries should begin with SELECT ALL ...; the set of columns from which data are returned is implicitly chosen by the choice of VAMDC-XSAMS format. If such a query does specify a set of columns (e.g. SELECT AtomIonCharge WHERE ...), then the query processor should ignore that set and proceed as if the query were SELECT ALL. However, where the results of a query are to be returned in a tabular format, the query processor must respect the query's selection of columns. In the latter case, if the query specifies a returnable not supported by the particular database then the processor should reject the query.

When processing a query that contains valid VSS1 plus extensions, the behaviour is defined by the implementation of the query processor. The processor may reject the query, or it may ignore the extensions that it does not support.

The following parts of SQL92 constitute VSS1 extensions: EXISTS, GROUP BY, HAVING, UNION, INTERSECT, EXCEPT, MINUS, ORDER BY, LIMIT, DECLARE, FETCH, CLOSE.

1.3 SQL92 syntax (informative)

The SQL92 standard [SQL92] should be consulted for the normative rules of syntax. These notes are for easy reference. VSS1 excludes so much of SQL that only the low-level aspects of the syntax are relevant.

SQL queries are written as text strings containing keywords, operators and operands separated by white space. Operands are names of tables and columns, sometimes called SQL identifiers or literal values. Identifiers and literals are sensitive to case; keywords and operators are not. There is a convention of writing keywords in upper case.

Queries can contain any Unicode character, but the keywords can be written using only ASCII characters. In VSS1, the valid identifiers also use only ASCII characters.

White space is required between keywords and operands but not between operators and operands. A typical (simple) VSS1 query looks like this:

```
SELECT ALL WHERE AtomIonCharge>6
```

This query would be equally valid:

```
SELECT * WHERE AtomIonCharge > 6
```

Here, data are selected from the columns AtomIonCharge and AtomNuclearCharge (note the use of a comma-separated list to specify the columns) of the table States according to a criterion on the electronic charge of the ions. String literals are delimited by single quotes (the ASCII apostrophe character) thus:

```
... WHERE AtomSymbol='Fe' ...
```

To include an apostrophe in a string, write two consecutive apostrophe-characters. If an identifier contains 'special characters' (typically white space), it must be protected with double quotes thus:

```
SELECT * WHERE "silly column name" > 0...
```

REFERENCES

[SQL92] Information Technology - Database Language SQL (Proposed revised text of DIS 9075) July 1992
<http://www.contrib.andrew.cmu.edu/~shadow/sql/sql1992.txt>