

# Data Type Mapping

## Contents

### Error rendering macro 'toc'

```
[com.ctc.wstx.exc.WstxLazyException] com.ctc.wstx.exc.WstxParsingException: String '--' not allowed in comment (missing '>?') at [row,col {unknown-source}]: [147,4]
```

HVR's mapping/conversion of data types is complex because each DBMS's data types have a specific range which seldom corresponds the range of another DBMS. For example, data type **varchar(10)** in SQL Server corresponds to **varchar2(10 bytes)** in Oracle, but **varchar(8000)** corresponds to **clob**. Note that the mapping here does not just depend on the 'name' of the data type, but also its 'attributes' like byte length, encoding, scale and precision, etc. If HVR is not accurate in mapping data types, then target tables could be created which are unable to contain the data delivered from the source DBMS.

Data types are not directly mapped/converted from the source DBMS to the target DBMS, instead they are first mapped to HVR's own 'catalog data types' and then these catalog data types are mapped to corresponding data types in target DBMS or file format. Data type mapping/conversion happens in the following manner:

1. When a channel is built using **Table Explore** or using the command **Hvradapt**, the source DBMS data types mapped to HVR's own 'catalog data types'. After the tables are added to a channel the mapping of data types in source DBMS (capture location) to HVR's catalog data types can be viewed in **HVRGUI** from the **Table Explore** window.
2. Then during **HVR Refresh** and **Integrate**, HVR's catalog data types are mapped to corresponding data types in the target DBMS (integrate location). This mapping happens at the moment the **HVR Refresh** is used to create target tables. After performing **HVR Refresh/Integrate** the mapping of HVR's catalog data types to the target DBMS data types can be viewed in **HVRGUI** from the **Table Explore** window.

## Customizing Data Type Mapping

If the automatic/default mapping of data type is not appropriate for a specific channel, it can be modified using action **ColumnProperties /Datatype**.

For example, by default HVR maps a **number** (without scale or precision) in Oracle to **numeric(38,4)** in SQL Server. By defining the following action, **number** (without scale or precision) in Oracle is mapped to **float** instead:

Group	Table	Action
SRCGRP	*	<b>ColumnProperties /DatatypeMatch="number[prec=0 &amp;&amp; scale=0]" /Datatype="float"</b>

In the above example, **ColumnProperties /DatatypeMatch** is used for mapping all columns with **number** (without scale or precision) data type into **float** data type.

Alternatively, **ColumnProperties /Name** can be used for mapping the data type of a specific column (e. g. **MYCOLUMN**) into **float** data type.

Group	Table	Action
-------	-------	--------

SRCGRP	*	<b>ColumnProperties</b> /Name="MYCOLUMN" /Datatype="float"
--------	---	--

### Extended Data Type Support

'Extended data types' are the DBMS data types (e.g. **sql\_variant** in SQL Server or **xmltype** in Oracle) which are not mapped to native HVR data types. Instead, HVR's **Extended Data Type Support** feature should be used for such data types.

HVR uses action **ColumnProperties /CaptureExpression** for converting an extended data type to a supported data type during **capture**, **compare** or **refresh** (source) and **ColumnProperties /IntegrateExpression** for converting it back to extended data type during **integrate**, **compare** or **refresh** (target) respectively.

### Data Type Mapping - Source to Target DBMS

In this section, you can find the mapping of data types in source DBMS to the corresponding data types in target DBMS or file format.

1. Select the required source DBMS in **Source Location Class**; displays the mapping of source DBMS data type to HVR catalog data type.
2. Select the required target DBMS in **Target Location Class**; displays the mapping of HVR catalog data type to target DBMS data type.

For example, to illustrate the data type mapping of **timestamp with time zone** in Oracle to the corresponding data type in Avro:

1. Select Oracle in **Source Location Class**.  
In the table displayed below **Source Location Class**, you can find that **timestamp with time zone** (in column **HVR Catalog Data Type**) is mapped to HVR's catalog data type **timestamp with tz (oracle)** (in column **HVR Catalog Data Type**).
2. Select Avro 1.6 and Avro 1.8 in **Target Location Class**.  
In the table displayed below **Target Location Class**, you can find that **timestamp with tz (oracle)** (in column **HVR Catalog Data Type**) is mapped to **varchar(100)** in Avro 1.6 and **timestamp-millis** in Avro 1.8.

#### Source Location Class

DB2 for i DB2 for Linux, UNIX and Windows  
DB2 for z/OS HANA Ingres MariaDB MySQL  
Oracle PostgreSQL SQL Server / Azure  
SQL Database

The version of the source location DBMS also affects how data types are mapped in HVR; the table below only corresponds to the latest DBMS versions.

#### Target Location Class

- The version of the target location DBMS also affects how HVR maps data types; It does not map to data types which are not supported by an older version of DBMS in target location. For simplicity, this is not reflected in the documentation here; the table below only corresponds to the latest supported DBMS versions.
- For data type Attributes (e.g. charlen, bytelen, scale, precision etc.) every feasible combination of values is possible. However, this table contains only enough combinations to illustrate the effect of variation.

Data Type	HVR Support	Capture HVR Data Type	HVR Catalog Data Type
bigint	Native		bigint
decfloat	Native		decfloat
decimal	Native		decimal
double	Native		double
integer	Native		integer
numeric	Native		numeric (db2i)
real	Native		real
smallint	Native		smallint
date	Native		ansidate
time	Native		time2

#### HVR Catalog Data Types Attributes

ansidate (ingres)

#### Aurora MySQL (UTF-8) Aurora PostgreSQL (UTF-8) Avro 1.6 Avro 1.8 Az Sy An

date date bigint date da

timestamp	Native	timestamp	ansidate		date	date	bigint	date	date
char	Native	char							
clob	Native	clob							
varchar	Native	varchar							
binary	Native	binary							
blob	Native	blob	bfile		longblob	bytea	blob	blob	varchar
varbinary	Native	varbinary							
datalink	Extended	<<datalink>>							
xml	Extended	<<xml>>							
rowid	Extended	<<rowid>>							
dbclob	Native	clob							
graphic	Native	nchar	bigint unsigned	bytelen=8	bigint unsigned	numeric(20)	float	decimal(20)	numeric(20)
vargraphic	Native	nvarchar							
user defined	Extended	<<user defined>>							
<b>Data Type</b>	<b>H V R Capture Support</b>	<b>HVR Catalog Data Type</b>	bigint	bytelen=8	bigint	bigint	bigint	bigint	bigint
bigint	Native	bigint							
decimal	Native	decimal	binary	bytelen=10	binary(10)	bytea	blob	blob	binary(10)
double	Native	double							
float	Native	double							
integer	Native	integer	binary_double	bytelen=8	double	float	float	float	float
numeric	Native	decimal							
real	Native	real							
smallint	Native	smallint	binary_float	bytelen=4	float	real	real	real	real
date	Native	ansidate							
time	Native	time							
timestamp	Native	timestamp							
char	Native	char							
char for bit data	Native	binary	bit (mysql)	bitlen=32	bit(32)	varchar(32)	varchar(32)	varchar(32)	varchar(32)
clob	Native	clob							
long varchar	Native	long varchar (db2)							
long varchar for bit data	Native	long varbinary							
varchar	Native	varchar							
varchar for bit data	Native	varbinary	bit		tinyint	boolean	boolean	boolean	bit
binary	Native	binary							
blob	Native	blob							
varbinary	Native	varbinary							
dbclob	Native	dbclob	blob		longblob	bytea	blob	blob	varchar
graphic	Native	graphic							
long vargraphic	Native	long nvarchar (db2)							
nclob	Native	dbclob							
vargraphic	Native	vargraphic							
xml	Native <sup>1</sup>	db2 xml <sup>1</sup>	boolean		tinyint	boolean	boolean	boolean	bit
rowid	Not supported								
user defined	Not supported								
<sup>1</sup> - Since HVR 5.6.5/2, the 'xml' data type is supported as 'Native' and is mapped to the HVR Catalog Type 'db2 xml'. Action <b>Capture /AugmentIncomplete=LOB</b> is required to capture xml type from DB2 for Linux, UNIX and Windows.									
<b>Data Type</b>	<b>H V R Capture Support</b>	<b>HVR Catalog Data Type</b>	byte varying	bytelen=10	varbinary(10)	bytea	blob	blob	varchar
smallint	Native	smallint							
integer	Native	integer							
bigint	Native	bigint							
decimal	Native	decimal	byteint	bytelen=1	tinyint	smallint	smallint	smallint	smallint
decimal	Native	decimal							
real	Native	real							
double	Native	double							
date	Native	ansidate							
time	Native	time2							













macaddr	Extended	<<macaddr>>								
macaddr8	Extended	<<macaddr8>>								
bit	Extended	<<bit>>								
bit varying / varbit	Extended	<<varbit>>	long char		encoding=UTF-8	longtext char		text	text	varchar
tsvector	Extended	<<tsvector>>								
tsquery	Extended	<<tsquery>>								
uuid	Native <sup>1</sup>	uniqueidentifier <sup>1</sup>								
xml	Extended	<<xml>>								
int4range	Extended	<<int4range>>								
int8range	Extended	<<int8range>>								
numrange	Extended	<<numrange>>								
tsrange	Extended	<<tsrange>>								
tstzrange	Extended	<<tstzrange>>	long nvarchar (db2)			longtext char t utf8mb4		text	text	nvarchar
daterange	Extended	<<daterange>>								
oid	Extended	<<oid>>								
pg_lsn	Extended	<<pg_lsn>>								
domain	Extended	<<domain>>								
user defined	Extended	<<user defined>>								

<sup>1</sup> - Since HVR 5.7.5/2 and 5.7.0/12, the 'uuid' data type is supported as 'Native' and is mapped to the HVR Catalog Type 'uniqueidentifier'.

Data Type	H V R Capture Support	HVR Catalog Data Type								
int	Native	int	long nvarchar			longtext char t utf8mb4		text	text	nvarchar
bigint	Native	bigint								
smallint	Native	smallint								
tinyint	Native	tinyint								
numeric	Native	numeric								
decimal	Native	numeric								
bit	Native	bit								
money	Native	money								
smallmoney	Native	smallmoney								
float	Native	float	long raw			longblob bytea		blob	blob	varchar
real	Native	real								
date	Native	ansidate								
datetime	Native	datetime								
datetime2	Native	datetime(2)								
datetimeoffset	Native	datetimeoffset								
smalldatetime	Native	smalldatetime								
time	Native	time								
char	Native	char	long varbinary			longblob bytea		blob	blob	varchar
varchar	Native	varchar								
text	Native	text (sqlserver)								
nchar	Native	nchar								
nvarchar	Native	nvarchar								
ntext	Native	ntext								
binary	Native	binary								
varbinary	Native	varbinary	long (db2) varchar		encoding=UTF-8	longtext char t utf8mb4		text	text	varchar
image	Native	image								
uniqueidentifier	Native	uniqueidentifier								
xml	Native	xml								
cursor	Extended	<<cursor>>								
hierarchyid	Extended	<<hierarchyid>>								
rowversion	Extended	<<rowversion>>								
sql_variant	Extended	<<sql_variant>>	long varchar		encoding=UTF-8	longtext char t utf8mb4		text	text	varchar
spatial geometry types	Extended	<<geometry>>								
spatial geography types	Extended	<<geography>>								
table	Extended	<<table>>	long		encoding=UTF-8	longtext char t utf8mb4		text	text	varchar

mediumint unsigned	bytelen=3	mediumint unsigned	integer	integer	integer	integer
mediumint	bytelen=3	mediumint	integer	integer	integer	integer
money (ingres)		decimal (14,2)	numeric (14,2)	varchar (16)	decimal (14,2)	money
money		decimal (19,4)	numeric (19,4)	varchar (21)	decimal (19,4)	money
nchar	charlen=10	char (10) charset utf8mb4	char(10)	char (10)	char (10)	nchar (10)
nchar (oracle)	charlen=10	char (10) charset utf8mb4	char(10)	char (10)	char (10)	nchar (10)
nclob		longtext charset utf8mb4	text	text	text	nvarchar (max)
ntext		longtext charset utf8mb4	text	text	text	nvarchar (max)
number		decimal (65,4)	numeric	float	decimal (100,4)	numeric (38,10)
number	prec=10scale=-127	double	float	float	float	float

number	prec=10scale=3	decimal(10,3)	numeric(10,3)	float	decimal(10,3)	numeric(10,3)
number	prec=26	decimal(26)	numeric(26)	float	decimal(26)	numeric(26)
number	prec=6	int	numeric(6)	integer	integer	int
numeric (db2i)	prec=10scale=3	decimal(10,3)	numeric(10,3)	varchar(12)	decimal(10,3)	numeric(10,3)
numeric (db2i)	prec=6	decimal(6)	numeric(6)	varchar(8)	decimal(6)	numeric(6)
numeric	prec=10scale=3	decimal(10,3)	numeric(10,3)	varchar(12)	decimal(10,3)	numeric(10,3)
numeric	prec=26	decimal(26)	numeric(26)	varchar(28)	decimal(26)	numeric(26)
numeric	prec=6	decimal(6)	numeric(6)	varchar(8)	decimal(6)	numeric(6)
nvarchar	charlen=10	varchar(10) charset utf8mb4	varchar(10)	varchar(10)	varchar(10)	nvarchar(10)
nvarchar(max)		longtext charset utf8mb4	text	text	text	nvarchar(max)

nvarchar2	charlen=10	varchar(10) charset utf8mb4	varchar(10)	varchar(10)	varchar(10)	nvarchar(10)
postgres date		date	date	bigint	date	date
postgres timestamp with time zone		datetime(0)	timestamp(0) with time zone	varchar(100)	timestamp-millis	datetime offset
postgres timestamp		datetime(0)	timestamp(0)	varchar(100)	timestamp-millis	datetime(2)
raw	bytelen=10	varbinary(10)	bytea	blob	blob	varbinary(10)
real	bytelen=4	float	real	real	real	real
rowid		char(18) charset ascii	char(18)	char(18)	char(18)	char
rowversion	bytelen=10	binary(10)	bytea	blob	blob	binary(10)
smalldatetime		datetime(0)	timestamp(0)	varchar(100)	timestamp-millis	datetime
smallint unsigned	bytelen=2	smallint unsigned	integer	integer	integer	int
smallint	bytelen=2	smallint	smallint	smallint	smallint	smallint

smallmoney		decimal(10,4)	numeric(10,4)	varchar(12)	decimal(10,4)	smallint
text (ingres)	bytelen=10encoding=UTF-8	varchar(10)charset utf8mb4	varchar(10)	varchar(10)	varchar(10)	varchar(10)
text (sqlserver)	encoding=WINDOWS-1252	longtext charset latin1	text	text	text	varchar(max)
text(sybase)	encoding=UTF-8	longtext charset utf8mb4	text	text	text	varchar(max)
time (mysql)	prec_sec=0	time(0)	time(0)	integer	time(3)	time
time (sybase)	prec_sec=0	time(3)	time(3)	integer	time(3)	time
time with local time zone		time(0)	time(0)	integer	time(3)	time
time with time zone		time(0)	time(0) with time zone	integer	time(3)	time
time	prec_sec=0	time(0)	time(0)	integer	time(3)	time

time	prec_sec=3	time(3)	time(3)	integer	time(3)	time
time2		time(6)	time(0)	integer	time(3)	time
timestamp (db2)	prec_sec=0	datetime(0)	timestamp(0)	varchar(100)	timestamp	datetime(2)
timestamp (ingres)		datetime(0)	timestamp(0)	varchar(100)	timestamp	datetime(2)
timestamp (mysql)		timestamp(0) null	timestamp(0)	bigint	timestamp	datetime(2)
timestamp (oracle)		datetime(0)	timestamp(0)	varchar(100)	timestamp	datetime(2)
timestamp (sqlserver)	bytelen=10	binary(10)	bytea	blob	blob	binary(1)
timestamp (sybase)	bytelen=10	varbinary(10)	bytea	blob	blob	varchar
timestamp with local time zone		datetime(0)	timestamp(0)	varchar(100)	timestamp	datetime(2)

timestamp with local tz (oracle)		datetime(0)	timestamp(0)	varchar(100)	timestamp-millis	date(2)
timestamp with time zone		datetime(0)	timestamp(0) with time zone	varchar(100)	timestamp-millis	date offset
timestamp with tz (oracle)		datetime(0)	timestamp(0) with time zone	varchar(100)	timestamp-millis	date offset
timestamp	prec_sec=0	datetime(0)	timestamp(0)	varchar(100)	timestamp-millis	date(2)
tinyint signed	bytelen=1	tinyint	smallint	smallint	smallint	signed
tinyint unsigned	bytelen=1	tinyint unsigned	smallint	smallint	smallint	tinyint
tinyint	bytelen=1	tinyint unsigned	smallint	smallint	smallint	tinyint
uniqueidentifier		binary(16)	uuid	char(36)	char(36)	uniqueidentifier
unitext		longtext charset utf8mb4	text	text	text	nvarchar(max)

univarchar	charlen=10	varchar(10) charset utf8mb4	varchar(10)	varchar(10)	varchar(10)	nvarchar(10)
unsigned bigint	bytelen=8	bigint unsigned	numeric(20)	float	decimal(20)	numeric(20)
unsigned int	bytelen=4	int unsigned	bigint	bigint	bigint	bigint
unsigned smallint	bytelen=2	smallint unsigned	integer	integer	integer	integer
urowid	bytelen=100	varchar(100) charset ascii	varchar(100)	varchar(100)	varchar(100)	varchar(100)
varbinary	bytelen=10	varbinary(10)	bytea	blob	blob	varbinary(10)
varbinary(sybase)	bytelen=10	varbinary(10)	bytea	blob	blob	varbinary(10)
varbinary(max)		longblob	bytea	blob	blob	varbinary(max)
varbyte	bytelen=10	varbinary(10)	bytea	blob	blob	varbinary(10)
varchar	bytelen=10 encoding=UTF-8	varchar(10) charset utf8mb4	varchar(10)	varchar(10)	varchar(10)	varchar(10)
varchar	bytelen=8000 encoding=UTF-8	varchar(8000) charset utf8mb4	varchar(8000)	varchar(8000)	varchar(8000)	varchar(8000)



varchar	bytelen=40 charlen=10 encoding=UTF-8	varchar(10) char(10) utf8mb4	varchar(10)	varchar(10)	varchar(10)	varchar(10)
varchar	bytelen=10 encoding=WINDO WS-1252	varchar(10) char(10) latin1	varchar(10)	varchar(10)	varchar(10)	varchar(10)
varchar (sybase)	bytelen=10 encoding=UTF-8	varchar(10) char(10) utf8mb4	varchar(10)	varchar(10)	varchar(10)	varchar(10)
varchar (sybase)	bytelen=8000 encoding=UTF-8	varchar(8000) char(8000) utf8mb4	varchar(8000)	varchar(8000)	varchar(8000)	varchar(8000)
varchar (sybase)	bytelen=40 charlen=10 encoding=UTF-8	varchar(10) char(10) utf8mb4	varchar(10)	varchar(10)	varchar(10)	varchar(10)
varchar (sybase)	bytelen=10 encoding=WINDO WS-1252	varchar(10) char(10) latin1	varchar(10)	varchar(10)	varchar(10)	varchar(10)
varchar(max)	encoding=WINDO WS-1252	longtext char(10) latin1	text	text	text	varchar(max)
varchar2	bytelen=10 encoding=UTF-8	varchar(10) char(10) utf8mb4	varchar(10)	varchar(10)	varchar(10)	varchar(10)

varchar2	bytelen=40 charlen=10 encoding=UTF-8	varchar(10) character set utf8mb4	varchar(10)	varchar(10)	varchar(10)	varchar(10)
varchar2	bytelen=10 encoding=WINDO WS-1252	varchar(10) character set latin1	varchar(10)	varchar(10)	varchar(10)	varchar(10)
vargraphic	charlen=10	varchar(10) character set utf8mb4	varchar(10)	varchar(10)	varchar(10)	nvarchar(10)
xml		longtext character set utf8mb4	text	text	text	nvarchar