Electronic data transfer format for Dakota County

Revision 02/25/2004 v6

Dakota County has adopted an electronic data transfer format defined by the Minnesota Department of Public Health (MDH). The format specifies 32 fields and the type of data in each field. The data may be submitted as a delimited ASCII text file (preferred), a .dbf file or Access Database file, an Excel spreadsheet, or as a fixed width ASCII text file. The format may be subject to change. Sample files are available from Dakota County at

http://www.dakotacounty.us

The data transfer table is named H20_XFER. This is a "flat file format" table, meaning that all applicable fields (columns) must be present in each record (row) of the file. Header lines and other comments or descriptions cannot occur in the defined fields (columns), and data import should not, in general, depend on such extra information. If stored in ASCII or spreadsheet form, the field data types might not be enforced by the software, and should be enforced by data entry personnel.

Several of the fields contain codes. All codes must be as defined below and in the sample tables, or should be documented in a separate file. Tables of codes and synonyms can be supplied by Dakota County via email if they are not already found in the sample files. Email your request to bill.olsen@co.dakota.mn.us

Several of the fields apply only to particular sampling programs or do not apply to every reported parameter. Fields may be left blank when they are inapplicable, but all fields should appear in the file. Fields marked as "Required" MUST contain valid data in every record (on every line).

Most Frequent Mistakes and Frequently Asked Questions

At the end of this document are compilations of the most frequent mistakes made in filling out the H20_XFER table. PLEASE review the table of Most Frequent Mistakes *before* filling out the H20_XFER table. There is also a list of Frequently Asked Questions.

Field names and data types for table: **H2O_XFER**

Field	Field Name	Type	Width	Data Required by Dakota County
1	SAMPLEVENT	Character	9	If available or appropriate
2	LAB_NO	Character	9	Yes
3	SAMPLE_NO	Character	10	Yes
4	RELATE_ID	Character	10	Yes (MN Unique Well Number)
5	CHEM_NO	Character	10	If available
6	CHEM_NAME	Character	26	Yes
7	AN_DATE	Date	8	If available
8	AN_METHOD	Character	10	If available
9	DETECTCODE	Character	2	Yes
10	RPT_LIMIT	Numeric	15.8	Yes
11	RESULT	Numeric	15.8	Yes
12	UNITS	Character	10	Yes
13	RESULT_UNC	Numeric	15.8	If available
14	LABCOMMENT	Character	10	If available
15	AGENCY	Character	7	If appropriate

16	PROGRAM_ID	Character	7	If appropriate
17	LIST_NO	Character	9	Not typically
18	FLD_SAMPNO	Character	10	If available
19	COLL_DATE	Date	8	Yes
20	COLL_TIME	Character	5	If available
21	COLL_NAME	Character	25	If available
22	T_BLANK	Character	10	If exists
23	M_BLANK	Character	10	If exists
24	F_BLANK	Character	10	If exists
25	RDS_FLAG	Logical	1	If appropriate
26	REASON	Character	2	If available
27	SEC_FLAG	Logical	1	If available
28	AN_RES_Q	Character	1	If available
29	FIELD_PROT	Logical	1	If available
30	LABQAQC	Logical	1	If available
31	RECVD_DATE	Date	8	If appropriate
32	REMARKS	Character	25	If appropriate

The Date (8) format in text is mmddyyyy. Any interpretable date format is acceptable in Access or Excell.

Field descriptions for table: **H2O_XFER**

SAMPLEVENT	Character	9	Client defined field for regularly scheduled sampling events or project identifier. Can be used for monitoring of regulatory compliance. (e.g. ”FQ1996” - Fall Quarterly Sampling for 1996)
LAB_NO	Character	9	Laboratory number of laboratory where sample was analyzed, If field analyzed, enter “FIELD”. Usually the laboratory certification number from the Minnesota Department of Health.
SAMPLE_NO	Character	10	Unique laboratory assigned sample number.
RELATE_ID	Character	10	Primary identifier for sample source or location. If the source was a well then the relate_id is the **Minnesota Unique Well Number.** ***Do NOT use local ID numbers such as “W1” or “MW32” here.*** If the sample is a Trip blank, Field blank, or Method blank, enter "T_BLANK", "F_BLANK", "M_BLANK", respectively. If the source of the analyte in the sample is a Spike, or Surrogate, enter "SPIKE", or "SURROGATE", respectively, rather than the source of the sample itself.
CHEM_NO	Character	10	Standard identifier for tested parameter from the **CHEM_NO** table. The Chemical Abstract Services (CAS) number, with dashes, is used as the CHEM_NO whenever available. *CHEM_NO codes can be assigned only by Dakota County or the MDH. If a CHEM_NO has not been assigned, leave this field blank, and ensure that the CHEM_NAME is unambiguous.*

CHEM_NAME	Character	26	An unambiguous name for the analyte or water quality parameter from the **CHEM_NO** table or **CHEM_SYN**onym table. *If the CHEM_NAME may be ambiguous, the CHEM_NO field should also be used.*
AN_DATE	Date	8	1.) For analytes which have sample information; the date on which the analysis was competed for this parameter. 2.) For analytes without sample information; the date that the sample was taken. Note that 'sample information' consists of those fields in this table following LABCOMMENT.
AN_METHOD	Character	10	Chemical analysis method from the **METHOD** table.
DETECTCODE	Character	2	Detection code for parameter, from **DETECTCODE** table. < less than NA Not Analyzed NQ Not Quantified - Negative = or Null value indicates that result field is non-zero. *If the* DETECTCODE *is "<", "=", or blank, then the* RESULT *field should have a numeric value entered, even if the value is zero.*
RPT_LIMIT	Numeric	15.8	1.) For analytes which have sampling information: enter the reporting limit for the parameter if it is different from the method detection limit. The default value is the method detection limit. 2.) For analytes without sampling information: the reporting limit is the same as the method detection limit and a null value indicates that no detection limit information is available for this result.
RESULT	Numeric	15.8	Result of the analysis. *This must be a number. If the result is, for example, "<1.0" then enter "<" in the* DETECTCODE *field, and "1.0" in the* RESULT *field. Do NOT enter "<1" in the* RESULT *field. For non-numeric results, such as "clear" or "smelly", enter the result in the* REMARKS *field*. For Spikes and Surrogates, enter the percent recovery, with "%" in the units field, or enter the actual measurement here, and the percent recovery in the REMARKS field, formatted like "108%".
UNITS	Character	10	Units of measure for RESULT and RESULT_UNC from **UNITS** table.
RESULT_UNC	Numeric	15.8	Result uncertainty, ±.
LABCOMMENT	Character	10	Laboratory comment for this result, from **LABCOMMENT** table

AGENCY	Character	7	Regulatory agency under which sample was taken, from **AGENCY** table. The purpose of this field is to allow finding of the original data, which is presumable archived by the Agency which has directed the data collection. *E.g., "MPCA" for sampling done under the VIC program.*
PROGRAM_ID	Character	7	Regulatory program under which sample was taken, from **PROGRAM_ID** table. *E.g. "VIC" for PCA VOLUNTARY INVESTIGATION & CLEAN UP*
LIST_NO	Character	9	MDH assigned identifier for the analyte list, laboratory, methods, and detection limits. ***The LIST_NO is normally left blank*** *if all of the following fields are filled in:* LAB_NO, SAMPLE_NO, CHEM_NAME, AN_METHOD, *and* RPT_LIMIT; *and if both detects and non-detects are reported.*
FLD_SAMPNO	Character	10	Field Sample Number. (A less preferred way to enter Blanks, Spikes and Surrogates is to enter "000000" in the RELATE_ID field, and the appropriate key word "T_BLANK", "M_BLANK", "F_BLANK", "SPIKE", or "SURROGATE" in this field.
COLL_DATE	Date	8	Sample collection date (The date format in text is MMDDYYYY) (Not required for Method Blanks)
COLL_TIME	Character	5	Sample collection time (The time format in text is HH:MM, where HH is on a 24 hour clock)
COLL_NAME	Character	25	Name of sample collector (preferred format; first initial, middle initial, and last name) from **WHO_IS** table, or program ID from PROGRAM_ID table, or MN Well contractor license no, or Lab Certification No.
T_BLANK	Character	10	Laboratory sample number of the Trip blank.
M_BLANK	Character	10	Laboratory sample number of the Method blank.
F_BLANK	Character	10	Laboratory sample number of the Field blank.

In records reporting results of a blank sample:

- Enter the *identical value* in the blank field as appears in the SAMPLE_NO field; this equality establishes that the result is of the blank. The type of blank is established by which of the blank fields is used. Entering "T_BLANK", "M_BLANK", or "F_BLANK",
- Enter "T_BLANK", "M_BLANK", or "F_BLANK" in the RELATE_ID field to confirm the blank type, unless the anlyte is a Spike or Surrogate. Do not leave the RELATE_ID field empty.

In records reporting results of well samples:

- Enter the lab sample number(s) of the blank(s) in the appropriate blank field(s); all three fields may be used. The results from the blank sample(s) are linked to the results of the well sample by equality of the lab sample number(s).
- If there is no blank associated with the well sample, then the

			respective blank field should be left empty. (See additional discussions below in the Most Common Errors Table and the FAQs section.)
RDS_FLAG	Logical	1	"T"/"F". Resample, Replicate, Duplicate, or Split sample flag. Associated sample numbers are entered in REMARKS. Type of sample is indicated in reason sample was taken, from the **REASON** table.
REASON	Character	2	Client defined* code for reason that sample was taken. * If the RDS_FLAG is set, enter one of the following codes from the **REASON** table: C, R, U, S to indicate Confirmation (resample), Replicate, dUplicate or Split, respectively.
SEC_FLAG	Logical	1	"T"/"F". Was the well secured before sample was taken?
AN_RES_Q	Character	1	Code for quality of the analytical results for sample from **AN_RES_Q** table. Also see County Well Index (CWI).: 1 - use with caution 2 - generally low reliability or accuracy 3 - normal reliability 4 - high reliability 5 - exceptionally high accuracy and precision
FIELD_PROT	Logical	1	"T"/"F". Was an approved field sampling protocol used when this sample was taken?
LABQAQC	Logical	1	"T"/"F". Does the laboratory have an approved QAQC plan on file for this sample?
RECVD_DATE	Date	8	Date on which the sample was received by the laboratory.
REMARKS	Character	25	Sample numbers of replicates and splits (see also the REASON field) Non-numeric results for analytes such as Odor or Color. Comments on field sample preparation procedure(s). Comments on analysis results. Percent recovery for Spikes and Surrogates, if actual values are reported in the RESULTS field.

Table of Most Frequent Mistakes

<table>
<tr><th>Field Name</th><th>Errors</th></tr>
<tr><td>RELATE_ID</td><td>Entry MUST be a Minnesota Unique Well Number, or one of the following key words:
T_BLANK, F_BLANK, M_BLANK, SPIKE, SURROGATE
Local Well IDs, such as MW-3, are not allowed.</td></tr>
<tr><td>SAMPLE_NO
RELATE_ID
FLD_SAMPNO
T_BLANK
M_BLANK
F_BLANK</td><td>If the sample is a Blank, Spike, or Surrogate, or if the sample is simply a well sample, but one or more Blank samples exist relating to this well sample, then the entries in these fields must satisfy certain relations. See the definitions of the fields above, and read the discussions in the Frequently Asked Questions, below.</td></tr>
<tr><td>DETECTCODE
RPT_LIMIT
RESULT</td><td>DETECTCODE must be one of the choices from the DETECTCODE table.
RPT_LIMIT and RESULT must be strictly numeric.
Here are some examples of correct and incorrect usage:
<table>
<tr><th></th><th>DETECTCODE</th><th>RPT_LIMIT</th><th>RESULT</th></tr>
<tr><td colspan="4">Example 1) The result is 98.0, which is above the reporting limit</td></tr>
<tr><td>Correct: “= 98.0”</td><td>=</td><td>10</td><td>98.0</td></tr>
<tr><td>Correct: a blank DETECT_CODE means “=”</td><td></td><td>10</td><td>98.0</td></tr>
<tr><td>INCORRECT: “=98.0” is not a number</td><td></td><td>10</td><td>=98.0</td></tr>
<tr><td colspan="4">Example 2) The result is a non-detect, or below the reporting limit</td></tr>
<tr><td>Correct: “< 10” means a non-detect</td><td><</td><td>10</td><td>10</td></tr>
<tr><td>Correct: “= 0” means a non-detect</td><td>=</td><td>10</td><td>0</td></tr>
<tr><td>INCORRECT: “<10” is not a code</td><td><10</td><td>10</td><td></td></tr>
<tr><td>INCORRECT: “ND” is not a defined code</td><td>ND</td><td>10</td><td></td></tr>
<tr><td>INCORRECT: “<10” is not a number</td><td></td><td>10</td><td><10</td></tr>
<tr><td>INCORRECT: “ND” is not a number</td><td></td><td>10</td><td>ND</td></tr>
</table></td></tr>
<tr><td>RESULT
REMARKS</td><td>Enter non-numeric type results in the REMARKS field
<table>
<tr><th></th><th>ANALYTE</th><th>DETECTCODE</th><th>RPT_LIMIT</th><th>RESULT</th><th>REMARKS</th></tr>
<tr><td>Correct: RPT_LIMIT and RESULT fields are left blank.</td><td>Color</td><td>=</td><td></td><td></td><td>Clear</td></tr>
<tr><td>INCORRECT: “Clear” is not a number</td><td>Color</td><td>=</td><td></td><td>Clear</td><td></td></tr>
<tr><td>INCORRECT: “Clear” is not a defined code</td><td>Color</td><td>Clear</td><td></td><td></td><td></td></tr>
<tr><td>INCORRECT: “NA” and “— “ are not numbers. Leave the RPT_LIMIT and RESULT fields blank.</td><td>Color</td><td>=</td><td>NA</td><td>--</td><td>Clear</td></tr>
</table></td></tr>
<tr><td>CHEM_NO
CHEM_NAME</td><td>The CHEM_NO is welcome but not required if the CHEM_NAME appears uniquely in the CHEMICAL SYNONYM Table.
The CHEM_NO, if known, or the CAS Number, if available (with dashes), must be entered in the CHEM_NO field if the CHEM_NAME is not entered, or if the CHEM_NAME appears more than once in the CHEMICAL SYNONYM table.</td></tr>
<tr><td>AN_DATE
COLL_DATE
RECVD_DATE</td><td>Dates must be in a legible format.
The standard format is mmddyyyy. Other formats will be rejected if they are not readable.
Please enter the entire 4-digit year, and avoid using a 2-digit year</td></tr>
<tr><td>Missing Data</td><td>Fields required by Dakota County for permit approval are defined on pages 1 and 2 of this document.</td></tr>
</table>

Table 1*. Most Frequent Mistakes*

Frequently Asked Questions (FAQs), and Tips.

1. *How should I enter Logical values?*
   a. In ASCII text files, enter Logical values as “T”/“F”, “Y”/”N”, or “1”/”0”.
   b. In Excell spreadsheets, enter Logical values as “TRUE”/”FALSE” or “-1”/”0”.
   c. In Access databases, define the field data type as Yes/No.

2. *How do I enter information about Trip Blanks, Field Blanks and Method Blanks?*

   The purpose of the following rules is to allow results from well samples and their associated blanks to be properly identified and related to each other.

   - In the rows reporting results from blanks:
     - Enter the "F_BLANK", "M_BLANK", or "T_BLANK" in RELATE_ID.
     - Enter the blank's lab sample no. in SAMPLE_NO and in the appropriate #_BLANK field.
     - Enter any associated blanks' sample no's. in the appropriate #_BLANK field.

   - In the rows reporting actual well sample results, for which there are associated blanks:
     - Enter the well's Unique Well No. in RELATE_ID.
     - Enter the well's lab sample no. in SAMPLE_NO.
     - Enter the blank's lab sample number in the appropriate #_BLANK field. All three of the #_BLANK fields may be filled in one record.

   - In the rows reporting actual well sample results for which there are no associated blanks:
     - Leave all three #_BLANK fields empty.

   The following examples illustrate two cases:
   A) 1 well sample is associated with 1 trip blank and 1 field blank. There are no field sample numbers, and the FLD_SAMPNO field is used for descriptive identifiers.
   B) 2 well samples are associated with 1 trip blank, the second is also associated with a field blank. Field sample numbers are used.

Case	***Description of data rows***	**SAMPLE_NO**	**RELATE_ID**	**T_BLANK**	**F_BLANK**
A	*Results from trip blank*	**2002100**	**T_Blank**		
	Results from field blank	**2002101**	**F_Blank**		
	Results from MW-1	**2002102**	**550001**	**2002100**	**2002101**
B	*Results from trip blank*	**2002210**	**T_Blank**		
	Results from field blank	**2002211**	**F_Blank**		
	Results from MW-2	**2002212**	**550002**	**2002210**	
	Results from MW-3	**2002213**	**550003**	**2002210**	**2002211**

Table 2. *Affected fields in the 2 example cases showing how to enter "blanks".*

Note that a single method blank sample number may be applied to results for several chemicals when different solutions are used for the different chemicals. Also note that the COLLECTION_DATE is not required for Method blanks. It is allowed to enter the date that the particular solution was mixed, or simply the date of analysis.

3. How do I enter results for Spikes and Surrogates?

A surrogate is a pure analyte added to an existing sample in the laboratory just before processing so that the overall efficiency of a method can be determined. The term Spike is usually used for an analyte that is already present in the sample, and the term Surrogate for an analyte not already present in the sample. They are treated identically here. (They differ from a Method blank, which consists of the pure analyte(s) in their own sample(s), though a Spike or Surrogate can be added to a Method blank.)

The results for Spikes and Surrogates are typically reported as the percent of the added analyte that is measured, called "percent recovery" (%). For example, if a surrogate is added to the sample to make a known concentration of 10.0 ug/L, and the lab measurement is 8.6 ug/L, then the surrogate result is 86% recovery. A result of 100% is ideal. Reports much different from 100% indicate miscallibration or matrix interference, and provide important information to the data user.

A surrogate has the same laboratory sample number as the sample it is added to, while a method blank has its own laboratory sample number. The fact that the analyte is added to the sample, and thus has a different origin than the sample, is made clear by entering "SURROGATE" in the RELATE_ID field.

A Spike or Surrogate result should be entered in its own row, and it should receive the same sample number as the sample to which it is added. The fact that the analyte is added to the sample, and thus has a different origin than the sample, is made clear by entering one of the key words "SPIKE" or "SURROGATE" in the RELATE_ID field.

The CHEM_NAME and CHEM_NO used for the surrogate analyte should be drawn from the **CHEMICAL** or **CHEM_SYN**onym table if possible. Do not modify the name of the analyte to show that it is a surrogate – i.e. don't change "TOLUENE" to "TOLUENE (S)". However, note that duterated compounds used for Surrogates have distinct Chemical names and numbers.

Enter the percent recovery in the RESULT field, and enter "%" in the UNITS field. The other results fields DETECT_CODE, RPT_LIMIT, LAB_COMMENT, etc. can entered as usual if the values are known, otherwise may be left blank. Enter the F_BLANK, M_BLANK, and T_BLANK fields as usual.

In the following 2 examples, see how the surrogate results are tied to the rest of the sample results by the SAMPLE_NO.

Description of data rows	**SAMPLE_NO**	**RELATE_ID**	**CHEM_NO**	**CHEM_NAME**	**RESULT**	**UNITS**	
Sample	**1001**	**523456**	**108-88-3**	***TOLUENE***	**13.1**	**ug/L**	
Spike in Sample	**1001**	**SPIKE**	**2037-26-5**	**TOLUENE-D8**	**86**	**%**	

Table 3a. *Example of well sample with spike (other fields are omitted for clarity).*

Description of data rows	**SAMPLE_NO**	**RELATE_ID**	**CHEM_NO**	**CHEM_NAME**	**RESULT**	**UNITS**	**T_BLANK**
Sample	**1002**	**500123**	**108-88-3**	***TOLUENE***	**13.1**	**ug/L**	**1003**
Spike in Sample	**1002**	**SPIKE**	**2037-26-5**	**TOLUENE-D8**	**86**	**%**	**1003**
Trip Blank	**1003**	**T_BLANK**	**108-88-3**	***TOLUENE***	**0**	**ug/L**	**1003**
Spike in Trip Blank	**1003**	**SPIKE**	**2037-26-5**	**TOLUENE-D8**	**107**	**%**	**1003**

Table 3b. *Example of well sample and trip blank for well 500123, both with spikes*

4. Where should I enter the local well identifier (e.g. “MW-1”)?

   The local well identifier is not required. However, entering it in the FIELD_SAMPNO field is a good way prevent data entry errors during manual data entry, and may help in discovering mis-identified wells. Make sure that you also enter the Minnesota Unique Well Number in the RELATE_ID field.

5. Where do I enter static water levels?

   Don't enter static water levels in the H20XFER table. Dakota County does not currently require static water levels to be submitted electronically, but you may submit them electronically, like the chemistry data, if you wish. *Please enter static water levels in a separate table*. A model table is not yet designed, but the minimum requirements are:

   Unique Well Number
   Date of measurement
   Water level
   Datum of the water level measurement

   No matter what form the data is submitted in, Dakota County requires that the datum of the measurement be stated. The datum can be in the form of a statement such as “feet below grade level”, “feet below top of casing”, “feet elevation”, “feet above local datum”, etc. If a local datum is used, Dakota County requires that you provide the elevation of the local datum. Dakota County does not prescribe a standard of precision for elevations or datum measurements: the datum may be as given by an actual survey (if avaialable, that is preferred), or may be interpreted from published contour maps.

6. How do I enter the results for Calibration Blanks?

   Calibration blanks are identified and entered as Method Blanks.

7. How can a lab enter the Minnesota Unique Well Number when the number is not provided to the lab?

   When results are reported to Dakota County under the Registered Well program, the results must be identified by the Minnesota Unique Well Number. Occasionally, a client does not provide the Minnesota Unique Well Number to the lab, but still wishes for the lab to prepare the data report for Dakota County. We suggest that the client may provide the Unique Well Number(s) to the lab after the analysis is reported, or the client may insert the Unique Well Number(s) in the data file before the file is submitted.

8. How do I enter Duplicates, Replicates, Splits, and Resamples?

   These samples are entered just as any other sample, with the addition of the information in the 3 fields shown below. The main thing about these samples is that they are associated with another sample; the association is carried in the REMARKS field. Fill in the following fields as shown:

RDS_FLAG	Enter "True" or "Y".
REASON	Enter one of the codes from the **REASON** table: C, R, U, S to indicate resample (Confirmation), Replicate, dUplicate or Split, respectively.
REMARKS	Enter the sample number of the associated sample.

9. How do I enter field measured parameters?

   Enter "FIELD" as the LAB_ID.
   Be sure to enter one of the analysis methods from the **METHOD** table.