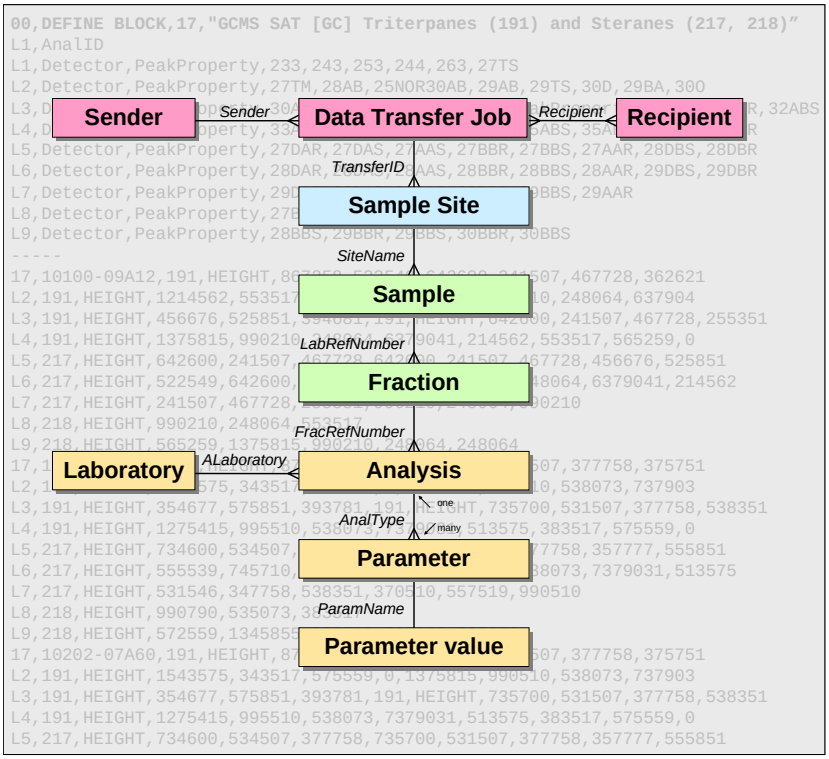
GC-NPD-95 format
for transfer of geochemical data

Version 2.0, May 2005



NORWEGIAN PETROLEUM DIRECTORATE

(This page is intentionally left blank.)

CONTENTS

PREFACE **5**

1. INTRODUCTION **7**
1.1 The complex structure of organic geochemical data 7
1.2 Purpose and basic structure of the GC-NPD-95 format 11

2. LOGICAL DATA MODEL **14**
2.1 Overview 14
2.2 Entity types and relations between them 16
2.2.1 Data Transfer Job 16
2.2.2 Sample Site 16
2.2.3 Sample 17
2.2.4 Fraction 18
2.2.5 Analysis 18
2.2.6 Parameter 19

3. DATA DICTIONARY **20**
3.1 General information 20
3.2 Remarks on individual attributes and their legal values 20
3.2.1 General remark 20
3.2.2 Legal attribute values not defined elsewhere 20
3.2.3 Legal dictionary names 20
3.2.4 Legal company codes 20
3.2.5 Legal country codes 21
3.2.6 Legal sample site names 21
3.2.7 Legal site type codes 21
3.2.8 Legal sample type codes 22
3.2.9 Legal lithology descriptors 22
3.2.10 Legal fraction type codes 23
3.2.11 Legal analysis type codes 23
3.2.12 Legal analysis description codes 23
3.2.13 Legal detector codes 23
3.2.14 Legal peak property codes 23
3.2.15 Legal parameter names 24

4. PHYSICAL DATA FORMAT **25**
4.1 Introduction 25
4.2 Transfer medium 25
4.3 General concepts and definitions 25
4.3.1 File type 25
4.3.2 File name 25
4.3.3 Character set 25
4.3.4 List delimiters, text encapsulation and record length 25
4.3.5 Decimal symbol and date format 26
4.3.6 Case sensitivity 26
4.3.7 Comment lines and blank records 26
4.3.8 Missing values 26
4.4 Format structure 26
4.5 Record formats 29
4.5.1 File definition lines 29
4.5.2 Block definition lines 30
4.5.3 Record definition lines 31
4.5.4 Value assignment records 35

5. MAJOR CHANGES SINCE VERSION 1.0 ..... **37**
5.1 Specification document 37
5.2 Logical data model 37
5.3 Data dictionary 38
5.4 Transfer media and physical format 38
5.5 Validation program 38

FIGURES

Figure 1 Relationships between samples, fractions, processes and results related to a geochemical analysis scheme. 8

Figure 2 General flow scheme of organic geochemical routine analyses (from NIGOGA). 10

Figure 3 Conceptual model for the transfer of geochemical data. 11

Figure 4 Basic structure of a GC-NPD-95 transfer file. 12

Figure 5 Schematic representation of the data transfer between two relational databases via a text (ASCII) file in GC-NPD-95 format. 13

Figure 6 The logical data model of GC-NPD-95. 15

Figure 7 Typical use of sample identifiers (LabRefNumber). 17

Figure 8 Encapsulation of text. All records shown are correctly encapsulated. 26

Figure 9 Structure of a GC-NPD-95 file. 27

Figure 10 Elements of a File definition block, a Record definition block and a Value assignment block. 28

Figure 11 Arrangement of data blocks in a GC-NPD-95 file. 29

Figure 12 General format of a file definition line. 30

Figure 13 Example of a file definition block. 30

Figure 14 General format of a block definition line (comma-delimited format). 30

Figure 15 General format of a record definition line (comma-delimited format). 31

Figure 16 A record definition block that defines one value assignment record format. 31

Figure 17 A record definition block defining three different record formats in the same block 32

Figure 18 Value assignment block containing parameter names. 32

Figure 19 A record definition block that declares parameter names using the explicit method. 33

Figure 20 A record definition block with declaration of parameter names using the implicit method. 33

Figure 21 A record definition block using the simplified implicit method. 33

Figure 22 The use of Detector and PeakProperty. 34

Figure 23 The use of Detector and PeakProperty in alternative format. 34

Figure 24 The use of PopnNum. 35

Figure 25 The use of PopnNum in alternative format. 35

Figure 26 Example of a value assignment record. 36

Figure 27 Value assignment record corresponding to the definition given in Figure 22 and Figure 23. 36

Figure 28 Value assignment records corresponding to the definition given in Figure 24 and Figure 25. 36

PREFACE

Character and content of this document

This document contains the technical specification of the format required by the Norwegian Petroleum Directorate (NPD) for the electronic transfer of (organic) geochemical data to the NPD. The requirement is stated in the "Reporting requirements for digital well data", a part of the Guidelines to the Petroleum Regulations which refers to Section 24 of the Resource Management Regulations[1,2].

The present document consists of this preface and five chapters:
1. A general introduction which describes the general structure of organic geochemical data and the principle of data transfer using a self-descriptive physical format.
2. A description of the *logical data model* used in the transfer files.
3. A description of the *data dictionary* with remarks on individual attributes.
4. A detailed description of the *physical data format*.
5. An overview of major differences between this Version 2 and the previous version of the data transfer format.

The data dictionary (GC-DIC-V2), which contains lists of legal entity types, attributes and attribute values, is contained in an Excel® workbook placed at NPD's website. This workbook can be downloaded for use with company-internal transfer programs.

A program for validation and format conversion of GC-NPD-95 files is also available at NPD's website, together with an annotated example of a transfer file.

Intended audience

This guideline is aimed toward database managers, software developers and geochemists who are concerned with electronic reporting of geochemical data according to the Resource Management Regulations. Chapter 1 (Introduction) attempts to explain the character and use of geochemical data to database managers and software developers, and introduces the idea of the self-descriptive format to geochemists. Chapters 2-5 are addressed to software developers who design and adjust transfer routines (programs) for digital geochemical data.

History

In 1990 the NPD adopted, with minimal modifications, Statoil's data transfer format for organic geochemical data as their "required format for geochemical data" (NPD document OD-90-98). This format defined fixed record formats for fixed selections of parameters.

When in 1993 the "Geochemical Standards Committee", consisting of representatives from Statoil, Saga, Norsk Hydro, IKU, Geolab Nor and the NPD, released the third edition of "The Norwegian Industry Guide to Organic Geochemical Analyses" (NIGOGA 3), the digital data

---

1 Regulations relating to resource management in the petroleum activities, 18 June 2001, Section 24. Final reporting of geological and reservoir technical well data. Norwegian Petroleum Directorate.

2 All regulations and guidelines mentioned in this document are available in digital form at NPD's website www.npd.no. (Search for "GC-NPD-95" or go to "Rules and Regulations"). The URLs to individual web pages may change and are hence not given in this document.

transfer format had to be adjusted and expanded. Members of the committee suggested that the old, rigid data format be replaced with a flexible, self-descriptive data transfer format that could easily be adapted to future needs.

The first version of the GC-NPD-95 standard data transfer format was developed in 1994-1995 by a work group consisting of representatives from the companies and institutes mentioned above. The format was based on the reporting requirements suggested by the NIGOGA 3. Version 1.0 of the GC-NPD-95 description was released in November 1995 as part of the “Provisions relating to digital transmission of geological and reservoir technical data in connection with the final report”, NPD document YA-061.

During the update in summer 2003 the format description was separated from the "Provisions" and was made accessible as a self-contained document in the World Wide Web.

The present Version 2.0 of GC-NPD-95 represents a major update. It considers some of the experiences made with the first version, and is adjusted to the most recently released edition of the NIGOGA. New and alternative attribute names and values were added and superfluous ones removed. The physical data format was changed into a column-independent, character-delimited format. The main differences between versions 1.0 and 2.0 are listed in Chapter 5 at the end of this document.

Relation between GC-NPD-95 and NIGOGA

The definitions contained in GC-NPD-95 are based on the rules and definitions given in the latest version of the NIGOGA. The NIGOGA also determines the units of measure associated with the individual parameters. The data dictionary, however, contains a wider range of attributes and attribute values than the NIGOGA requires (elemental composition, attributes for the characterisation of shallow sediment samples etc.).

Version names and responsibility for updates

The format specified in this document is named "GC-NPD-95 version 2.0". The NPD is responsible for releases of new formats.

The dictionary specified in this document is named "GC-DIC-V2". The Norwegian oil industry, by its appointed "NIGOGA committee", is responsible for the release of new dictionaries.

Contributors to the original version of GC-NPD-95 (version 1.0)

Data format and data dictionary: NPD, DLISAS (Derek Lacey), IKU (Hermann M. Weiss), Norsk Hydro (Trond Hanesand), Saga Petroleum (Nigel Mills), Statoil (Richard Patience).

Contributors to the present version (version 2.0)

Specification and data dictionary: NPD, Brekke Chemo (Trond Brekke), SINTEF Petroleum Research (Hermann M. Weiss). Validation program: Infospine (Børge Hansen, Dag Hagen). APT, Geolab Nor, IGI, Norsk Hydro, OLF, Statoil and SINTEF Petroleum Research gave feedback during the hearing round.

1. INTRODUCTION

1.1 The complex structure of organic geochemical data

People unfamiliar with geochemical analyses often ask why geochemical data should require a complex data transfer format. This section tries to give an explanation.

Sample Site information
Analytical data can only be meaningfully interpreted if one knows the exact origin of the sample material, i.e. the exact point in the earth's crust from which the sample was taken. Parts of the necessary information may be implicit from the location name, e.g. with exploration wells, where these details (coordinates, depth, datum etc.) are easily accessible in public databases. In other cases the full details are not publicly known and must therefore be fully specified when the corresponding analytical data are transferred.

Sample information
Organic geochemical analyses can be carried out on various kinds of sample materials (e.g. rock, oil, gas), which can be available in the form of different *sample types* (e.g. rock sample types like canned drill cuttings, cores or side wall cores, or the sample types *oil* and *gas*). The sample type often greatly controls the quality and reliability of the results and is therefore essential information for the interpreter.

Fraction information
Some analyses, e.g. the determination of the total carbon content, could theoretically be performed on the untreated whole sample (Figure 1a). By far most analyses - except for e.g. lithological description - will be carried out on *fractions*, i.e. preparations or specimens of the original sample that are considered “homogeneous” for the purpose of the analysis performed. These are obtained by separation (picking, sieving, filtering, precipitation etc.) and/or some preparation or mounting process (crushing, casting and polishing etc.)[3] (Figure 1b, c). A polished block of kerogen concentrate, or the saturated hydrocarbon fraction prepared from the solvent extract of a crushed rock sample are examples of fractions.

Fractionation of a sample may not only yield the fractions themselves, but also numerical results (e.g. precipitation of asphaltenes from oil yields two fractions and a result, the asphaltene content of the oil; Figure 1d). Fractionation can be carried out in several consecutive steps, leading to a hierarchy of "parent" and "child" fractions (Figure 1d).

It is crucial for the interpreter to know exactly on which fraction an analysis was performed, as an analytical method (e.g. gas chromatography) often can be applied to several different fractions (e.g. the whole oil, the saturated hydrocarbon fraction or the aromatic hydrocarbon fraction; Figure 2). These analyses may yield the same result parameters (e.g. concentration, peak height, isotope ratio etc. for a certain n-alkane), but these have, of course, different geochemical significance depending on the fraction analysed.

[3] In a narrow sense, a fraction is the product of separation only, but for practical reasons we include the type of preparation in the *fraction* term.

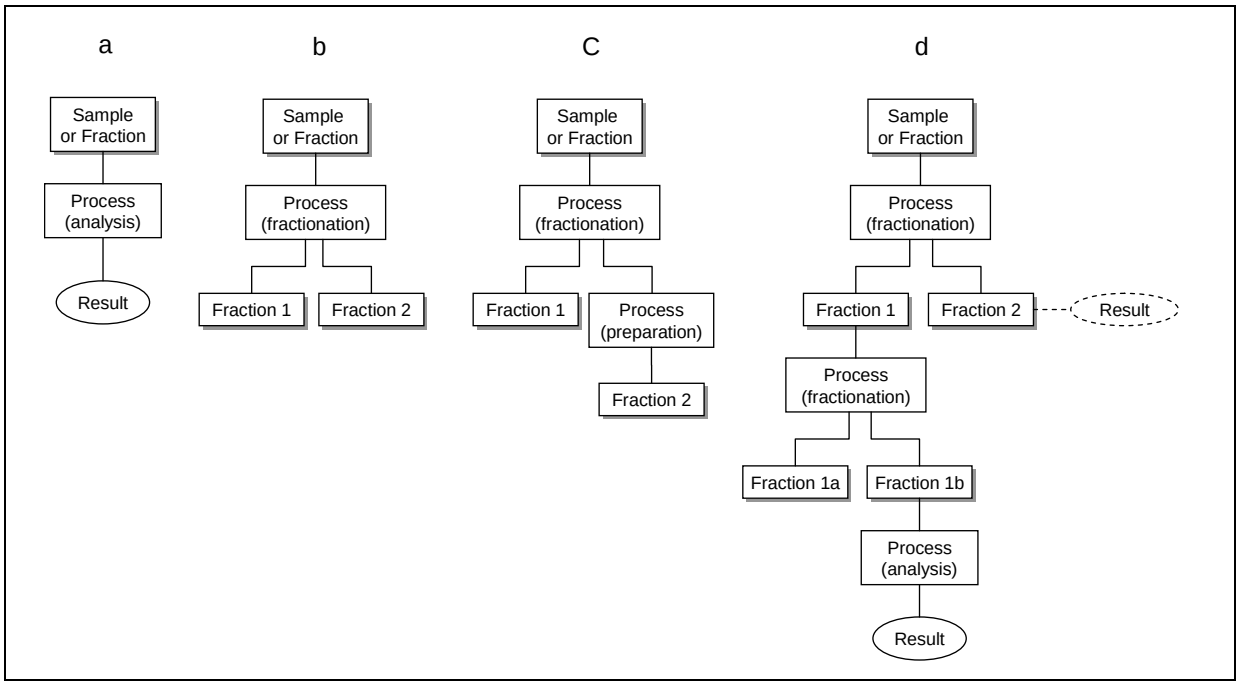


Figure 1 *Relationships between samples, fractions, processes and results related to a geochemical analysis scheme.*

Analysis information
A wide spectrum of analytical methods and analytical conditions is used in geochemical analysis, and exact information on this is indispensable for the user of the results.

Parameter information
Most analytical techniques yield not a single value but a number of values for different parameters. It is for example usual to quantify between 10 and 50 different peaks in a gas chromatogram or mass fragmentogram. These must be uniquely identified.

The abundance of a compound detected by gas chromatography or GC-MS[4] can be expressed in terms of peak height, peak area or concentration. Concentration can in turn be derived from either peak height or peak area; it can relate to mass, volume etc. and may be expressed in different units of measure. This leads to a large number of possibilities (and numerically different results), and different laboratories and oil companies may have different preferences.

The user of the data needs to know exactly what has been measured and in which unit the result is expressed.

Information on parallel/repeated analyses
A specimen prepared from a sample or fraction may be analysed repeatedly (if the analysis is non-destructive), or different aliquots (i.e. parts) of a (homogeneous) sample or fraction may be analysed at different times, at different laboratories, or by different analytical techniques.

4 gas chromatography-mass spectrometry

Information on control analyses
Reference materials may be analysed together with the samples for control and comparison between different sets of analysis results. The results from these control analyses must be clearly identifiable as such.

Additional information (comments, dates)
Finally, neither samples nor analyses are always perfect, and comments may be necessary to help the interpreter to understand the results or to assess their quality and relevance. Sampling or analysis dates may be important information, e.g. to identify compositional changes during production or well tests. The analysis date may also help to trace analysis conditions etc.

Because of all these possibilities, the actual analytical flow scheme may become quite complex (Figure 2), and the geochemist typically requires information on all processes and their sequence in order to interpret the data meaningfully. The data transfer format should therefore take care of as much as possible of this information in order to allow the reconstruction of the analytical flow, which can be generally expressed by a data model such as the one shown in Figure 3 (where also the data transfer itself is included).

(This space is intentionally left blank.)

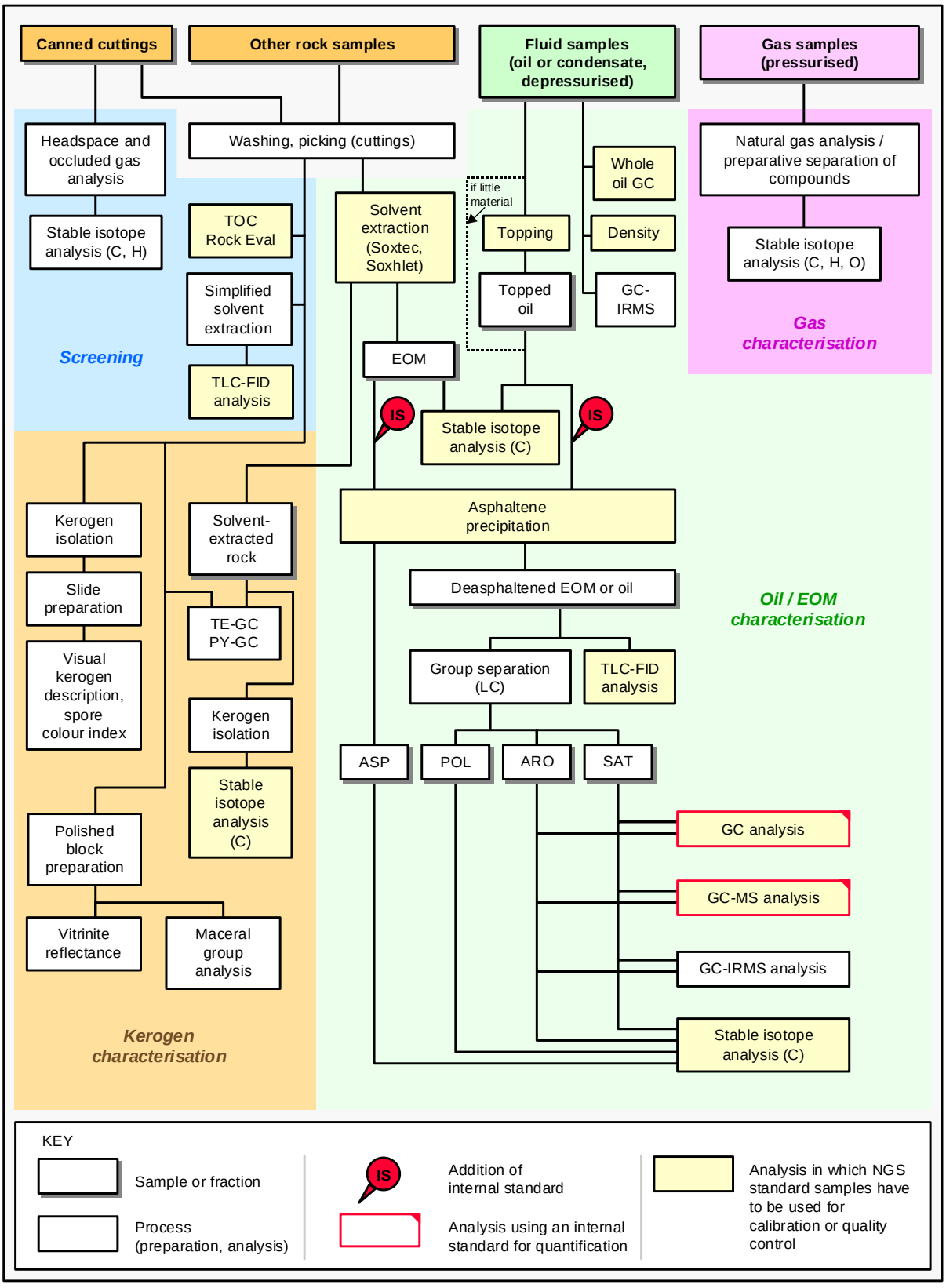


Figure 2 *General flow scheme of organic geochemical routine analyses (from NIGOGA). Note that the same analytical method can be applied to different fractions of the same sample.*

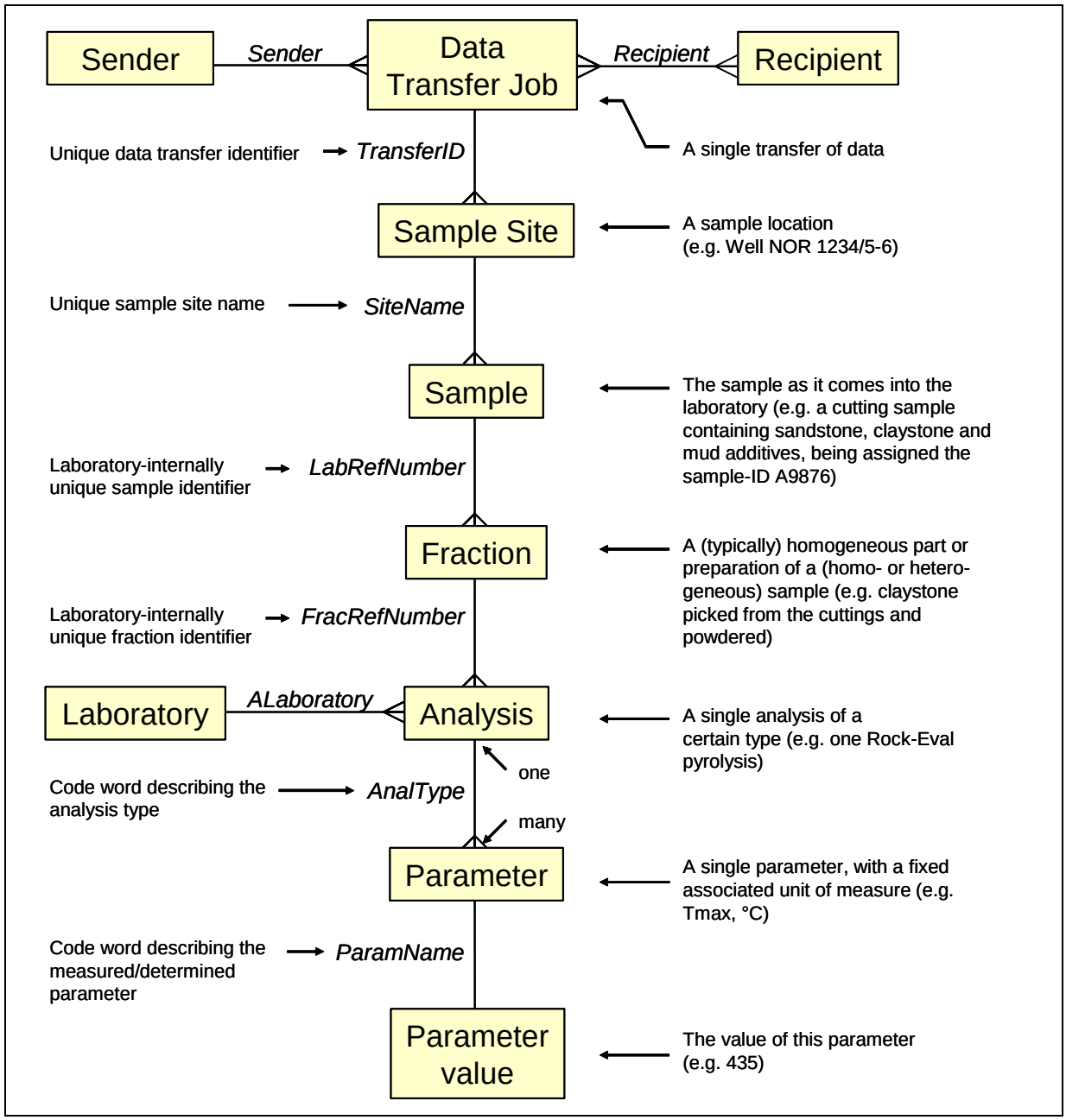


Figure 3 Conceptual model for the transfer of geochemical data.

1.2 Purpose and basic structure of the GC-NPD-95 format

The GC-NPD-95 format (GC stands for GeoChemistry, 95 for the year 1995) is a self-descriptive transfer format that intends to cater for the described variability and complexity of geochemical data and to allow an easy adjustment of the data transfer to changing or new analytical methods.

Please note that the GC-NPD-95 format is designed for the communication between computers. Files in this format are NOT a substitute of the “normal” tables (spreadsheets, PDF files etc.) which are intended to be read by humans.

The principle of GC-NPD-95 is as follows:

(1) The basic entity types (e.g. "sample"), their attributes (e.g. "lithology") and their relationships (e.g. several samples can belong to the same sample site, but not vice versa) are defined in a *logical* (or *conceptual*) *data model* which is known to all users (or data systems).

(2) All parameters (variables) to be transferred, the units of measure (UoM) associated with these parameters[5], and legal attribute values (e.g. sample types) are defined in a *data dictionary* which is known to all users (or data systems).

(3) The transfer file is organised in pairs of *record definition blocks* and *value assignment blocks*. A *record definition block* contains attribute names (and formats), and the corresponding *value assignment block* contains the values for these attributes, i.e. the actual results. The respective units of measure are implicit from the data dictionary and hence not included in the transfer file.

(4) The version of the data dictionary and the type and version of the transfer format are, together with information on the source and context of the data, contained in the *file definition block*, which forms the first few lines of the transfer file (Figure 4).

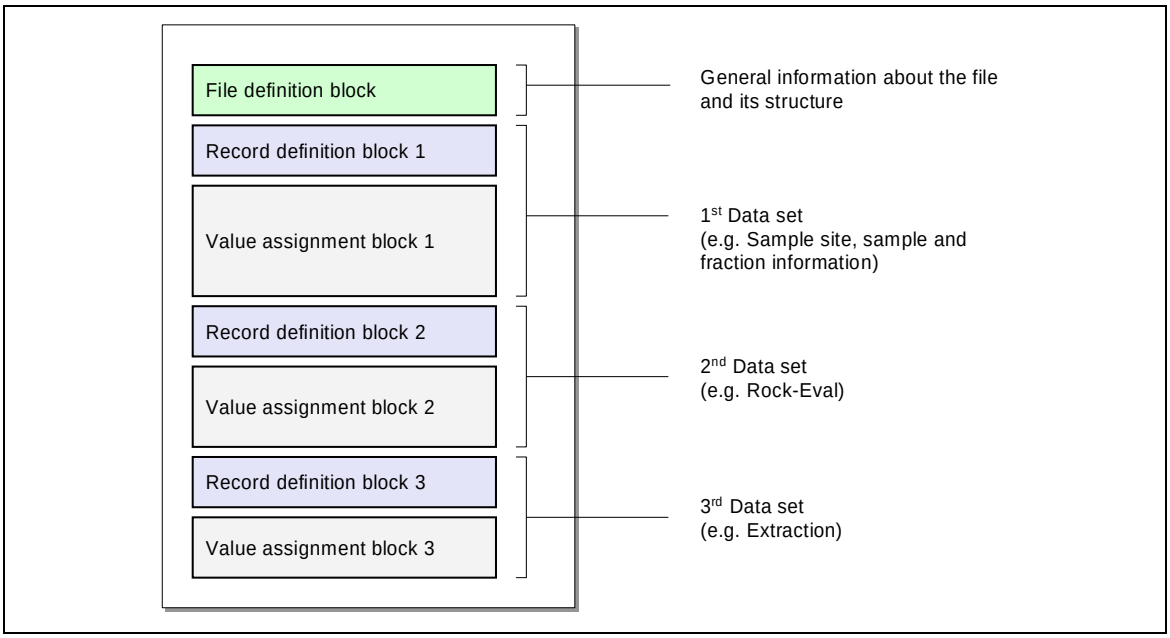


Figure 4 Basic structure of a GC-NPD-95 transfer file.

A single GC-NPD-95 file can contain several data sets, each having its own *record definition block* (Figure 4). Each data set may represent the results from a different type of analysis (e.g. Rock-Eval, extraction, SAT GC). It is also possible to break the results from one type of analysis (e.g. SAT GCMS) up into several data sets (e.g. one for each ion mass). The freedom in changing the sequence of these data sets and in defining the content and format of each of them, provides the flexibility desired in the transfer of complex data sets.

5 GC-NPD-95 allows only a single unit of measure for each parameter.

The reading program that "knows" the data structure and the dictionary can then recognise the physical structure of the transfer file, identify the parameters related to the reported values and write the values correctly to the receiver's database (Figure 5).

It is not necessary that two databases that exchange data via the GC-NPD-95 format have the same *physical* structure (i.e. the same tables and the same logical relationships between these tables). The only requirement is that the two database structures are compatible with the common *logical* data model and that the software used to transfer the data between them recognises a common data dictionary (Figure 5).

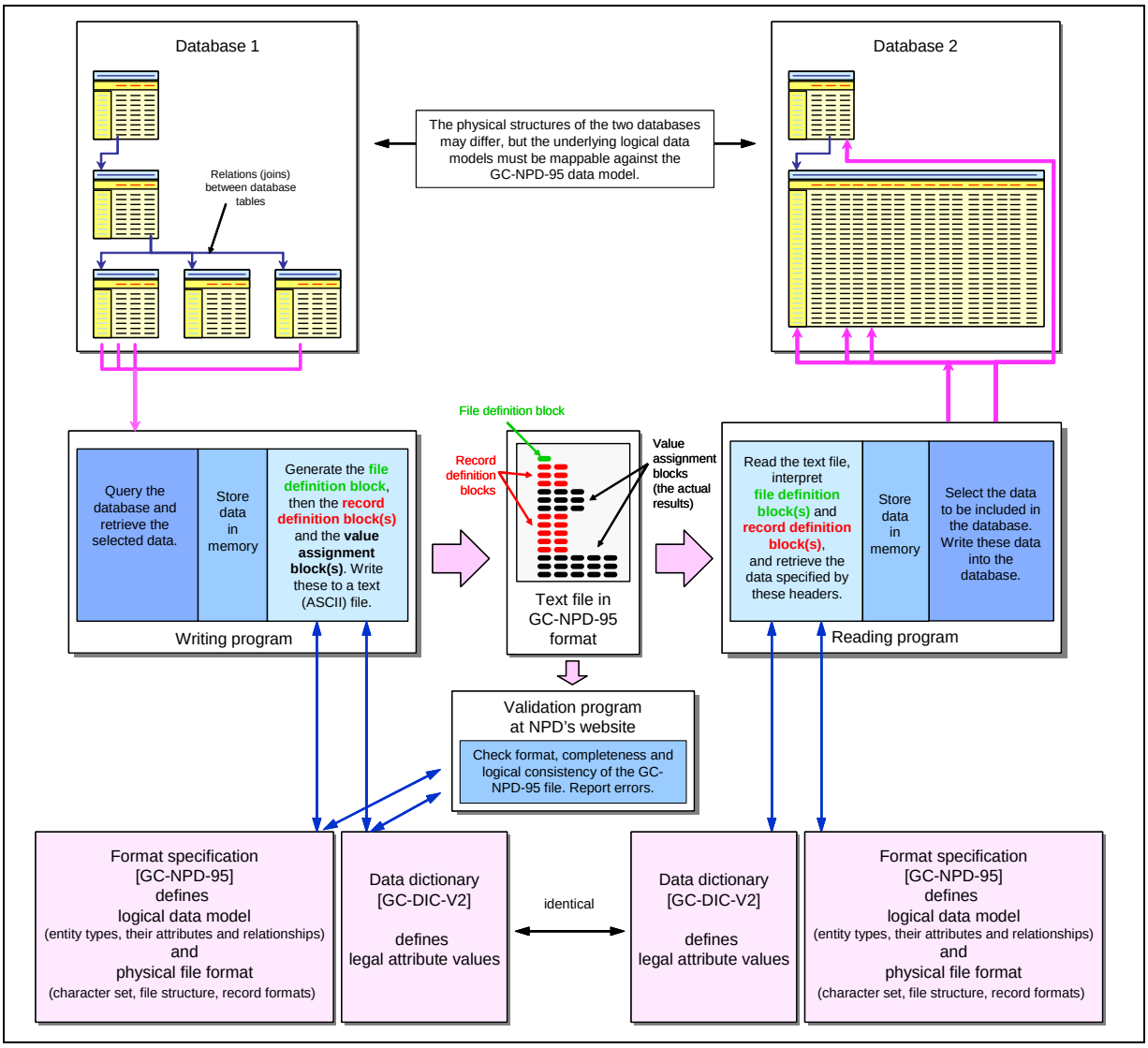


Figure 5 *Schematic representation of the data transfer between two relational databases via a text (ASCII) file in GC-NPD-95 format. Sender's database to the left, receiver's database to the right. Pink arrows: flow of analytical data, Blue arrows: Lookup of dictionary information. The validation program is also indicated.*

2. LOGICAL DATA MODEL

2.1 Overview

The logical data model used for the physical implementation of the GC-NPD-95 format is illustrated in Figure 6. It is based on the analytical procedures and reporting requirements defined in the NIGOGA.

The model contains six entity types (Data Transfer Job, Sample Site, Sample, Fraction, Analysis and Parameter) and is similar to the conceptual model shown in Figure 3. A major difference, however, is that in addition to the laboratory-internal keys (*SiteName, LabRefNumber, FracRefNumber* and *AnalType*[6]), transfer file-internal keys are introduced to make the transfer file structure completely self-contained (*SiteID, SampleID, FractionID, and AnalID*). This enables a proper transfer also if any of the laboratory-internal keys should be missing (e.g. with "old" data), or if the internal keys from different laboratories accidentally should be identical.

Each entity type possesses a number of attributes. For example, sample type (*SampleType*) and lithology (*SLithology*) are attributes of *Sample*, and an individual measurement result (*ParamValue*) is an attribute of the entity *Parameter*.

Some of the attributes are mandatory (i.e. must always be reported, red in Figure 6), some are conditional (i.e. they must be reported in certain circumstances, green in Figure 6) and others are optional (i.e. they do not necessarily have to be reported, blue in Figure 6). Attributes used to identify entities and link them logically together (*TransferID, SiteID, SampleID, FractionID, AnalID*) are always mandatory. The three conditional attributes of the *Parameter* entity (*PopnNum*, *Detector* and *PeakProperty*) are linked to specific analysis types that can produce multiple numerical results for the same parameter during a single analysis run (vitrinite reflectance and gas chromatography).

Some attributes can only assume certain values (e.g. certain country codes). These “legal” values are listed in dictionary tables (lookup tables, blue boxes in Figure 6)[7].

For historical (“old”) data, mandatory or conditional information may be unavailable. The value “UNKNOWN” is therefore accepted as legal value for certain mandatory or conditional data attributes. Note that this applies ONLY to historical data.

The entity types are defined in the following chapter. Attributes and legal attribute values are described in the GC-NPD-95 Data Dictionary “GC-DIC-V2”. This is an Excel® workbook available at NPD’s web site. The structure of this dictionary is described in Chapter 3 of this specification where also some comments on individual attributes are included.

---

6 Identical to *AnalName* in version 1.0.

7 The only exceptions are two attributes of *Data Transfer Job* whose legal values are defined below in this specification.

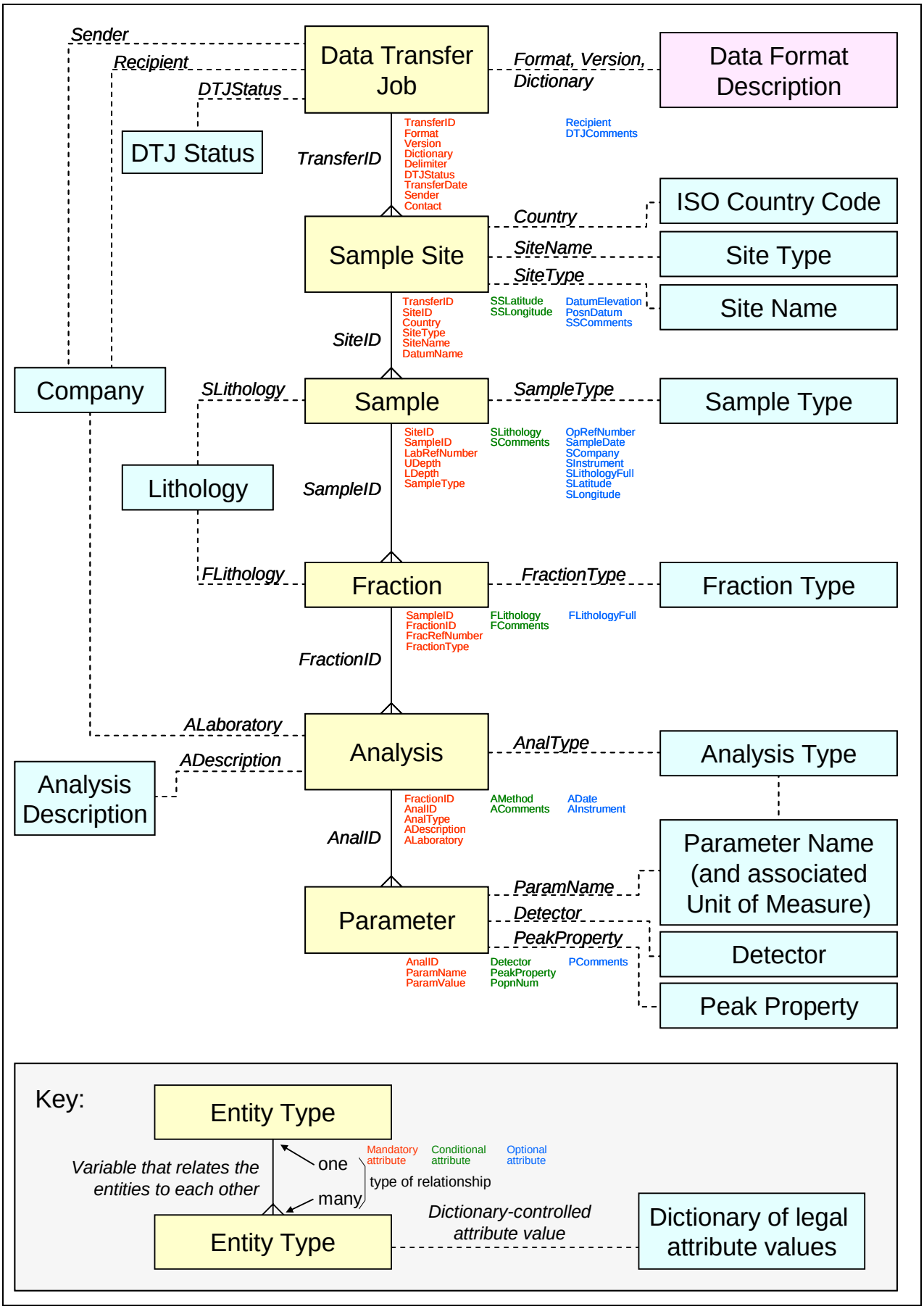


Figure 6 The logical data model of GC-NPD-95.

2.2 Entity types and relations between them

This section describes the six entity types of GC-NPD-95. The associated attributes are listed in the GC-DIC-V2 data dictionary (Excel® workbook).

2.2.1 Data Transfer Job

Definition

"The identification of a set of data that is being transferred."

Legal attribute values not defined elsewhere

Attribute	Explanation	Variable type	Legal values
Format	Name of data transfer format	text	GC-NPD-95
Version	Version of data transfer format	text	2.0

Remarks

- A Data Transfer Job corresponds to the transfer of exactly one physical transfer file.
- The attributes of the Data Transfer Job communicate administrative data concerning the data transfer, such as originating organisation, contact person, date and reference information, and any other information that is required to uniquely identify a data set and the format, version and dictionaries being referred to.

Relation to other entity types

The Data Transfer Job is the top-level entity type. One Data Transfer Job is logically related to one or several Sample Sites via the common attribute *TransferID*.

2.2.2 Sample Site

Definition

"The description of a location from which samples are collected."

Remarks

- The attributes of Sample Site allow the unambiguous identification and determination of any sample position. They cater for the reporting of data collected from a variety of location types, such as wells, geographical areas, outcrop locations or any other kind of site.

Relation to other entity types

A Sample Site is logically related to exactly one Data Transfer Job (via the common attribute *TransferID*) and to one or several Samples (via the common attribute *SiteID*).

2.2.3 Sample

Definition

"The description of the whole material as originally retrieved/sampled by the client or operator." (Or less formal: “A sample is the material as it comes through the door to the laboratory.”).

Relation to other entity types

A Sample is logically related to exactly one Sample Site (via the common attribute *SiteID*) and can contain one ore several Fractions (logically linked via the common attribute *SampleID*).

Remarks

- Downhole samples from deviated wells may have geographical locations that differ from that of the sample site (wellhead).
- Note that Lower and Upper depth shall always be reported, even if they are identical (“spot samples”). Lower depth is the depth normally referred to in tables and figures (NIGOGA).
- This model does not cater for samples from multiple, non-contiguous depth intervals of type OIL (DST), DCOM (composite sample) etc. With such samples use the bottom depth of the deepest and the top depth of the shallowest sample interval.
- *OpRefNumber*, the operator’s sample identifier, is mandatory for reference samples (standards). The NGS samples have the following *OpRefNumber*s: “NGS NSO-1”, “NGS SR-1” and “NGS JR-1”.
- *LabRefNumber*, the laboratory’s sample identifier, is mandatory and must be unique within each laboratory in order to trace a sample and its associated data. A sample may get several different *LabRefNumber*s, as each laboratory normally assigns its own identifier (*LabRefNumber*) to incoming samples. Sample identifiers provided by the customer (e.g. operator) are typically stored in the laboratory's database, but not used for lab-internal sample identification (Figure 7). The GC-NPD-95 format uses an additional transfer-file internal sample identifier (*SampleID*) to avoid any ambiguity and to enable proper transfer also if any of the laboratory-internal reference codes should be missing (e.g. with "old" data).

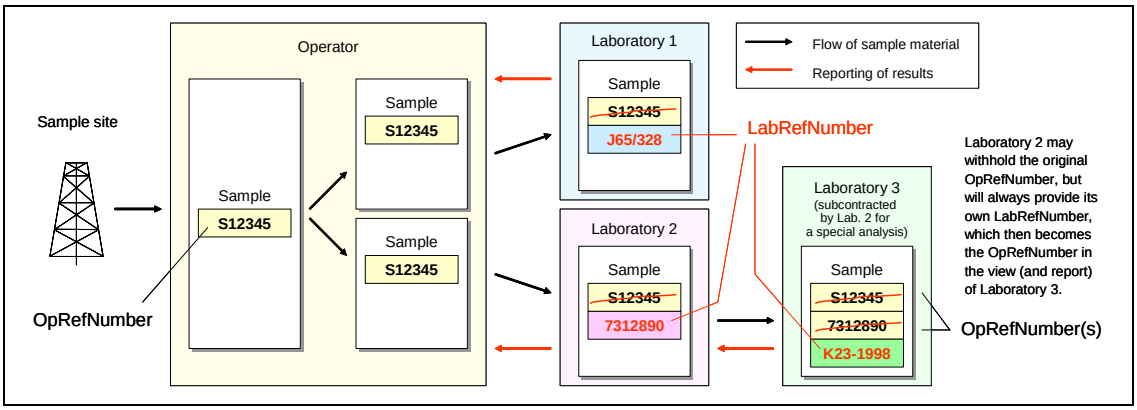


Figure 7 Typical use of sample identifiers (LabRefNumber).

2.2.4 Fraction

Definition

"The laboratory preparation of material from a sample, which is treated as "homogeneous" for analytical purposes."

Relation to other entity types

A Fraction is logically related to exactly one Sample (via the common attribute *SampleID*) and can be subjected to one ore several Analyses (logically linked via the common attribute *FractionID*).

Remarks

- The transfer of fraction data is mandatory, as this information is typically necessary for the interpretation of the results.
- *FractionType* describes the type of material and/or the type of preparation.
- The fraction types defined in GC-NPD-95 allow tracking of the typical fractions generated during routine well analyses (i.e. determination of the fraction itself and its parent fraction). The data model does not cater for more complicated analytical procedures where numerous fraction generations with complex relationships are generated (e.g. hydrous pyrolysis, HPLC separation of aromatics etc.). Non-standard fractions should be assigned the fraction type “Other” which then must be described using the *FComments* attribute.
- With rock samples, fractions (e.g. rock extract, *FractionType* EOM) may be derived from the bulk sample (which can contain various different lithologies) or from a picked lithology (which after crushing gets the *FractionType* SDUN). To preserve this information, the lithology of the fraction (*FLithology*) must be reported for all rock-derived fractions. Note that GC-NPD-95 contains no fraction type for “picked cuttings”, as these are never analysed without further preparation (crushing, casting into epoxy blocks etc.).

2.2.5 Analysis

Definition

"The description of the analysis performed."

Remarks

- The attributes allow to specify all relevant information concerning the conditions under which the analysis was run.
- Multiple analyses of the same fraction (e.g. of a homogeneous reference standard) will be assigned different *AnalID*s, which allows their unambiguous identification.

Relation to other entity types

Each instance of an Analysis is carried out on exactly one Fraction (logically linked via the common attribute *FractionID*) and can generate results/values of one or several Parameters (logically linked via the common attribute *AnalID*).

2.2.6 Parameter

Definition

”The individual parameters and their measured values.”

Remarks

- A set of parameters is defined for each analysis type.
- Parameter usage is, if necessary, explained in the parameter list of the data dictionary.
- Legal parameter names are specified in the *ParameterName* sheet of the data dictionary.
- Exactly one unit of measure (UOM) is implicitly associated with each parameter. This UOM is defined in the NIGOGA.

Relation to other entity types

A Parameter is logically related to exactly one Analysis (via the common attribute *AnalID*). *Parameter* represents the lowest level in the hierarchy of entity types.

3. DATA DICTIONARY

3.1 General information

The GC-NPD-95 Data Dictionary “GC-DIC-V2” is an Excel® workbook available at NPD’s web site. This workbook is the only official documentation of the definitions of attributes and legal attribute values used by the GC-NPD-95 V2.0 format. It also acts as the data dictionary for the validation program. This program can be run via the Internet at NPD's web site to read and check GC-NPD-95 (V2) files for correct formatting and internal logical consistency. The dictionary can be downloaded free of charge and shall be used as reference by all customer programs for generating or importing GC-NPD-95 transfer files.

The workbook contains a list of attributes (sheet “Attribute”) and various sheets containing legal attribute values defined in GC-NPD-95, reference codes to attributes etc. In addition it contains a sheet that defines the valid name of the data dictionary (“Dictionary”), a sheet that defines different validity tests carried out by the reading program (“CheckType”). Where necessary, the sheets contain definitions, explanations and remarks on the usage of the various attributes and attribute values. A “UserInformation” sheet gives an overview on the contents of the worksheets and explains the relationships between them.

3.2 Remarks on individual attributes and their legal values

3.2.1 General remark

Some values listed in the dictionary shall be applied only to "old" data, where some information cannot be reconstructed (e.g. UNKNOWN). Other values were included to allow reporting of non-standard entities, e.g. untypical sample or analysis types.

3.2.2 Legal attribute values not defined elsewhere

The attributes *Format* and *Version* of the entity type *Data_Transfer_Job* are not included in the dictionary as they have only one legal value each (see Section 2.2 of this specification).

3.2.3 Legal dictionary names

The legal value of *Dictionary* (an attribute of *Data_Transfer_Job*) is contained in the “Dictionary” sheet of the workbook.

3.2.4 Legal company codes

Legal company codes are listed in the “LookupValue” sheet of the workbook. They apply to *Sender* and *Recipient* (attributes of *Data_Transfer_Job*), *SCompany* (an attribute of *Sample*) and *ALaboratory* attribute of *Analysis*).

Remarks

- The table contains the codes and full names of oil companies and geochemical laboratories, including some defunct enterprises in order to allow the transfer of "old" data. The names were updated as far as possible, using official documentation and the companies' own Web pages. The official names of the enterprises, however, are changing

continually; hence this list will never be completely up to date. In cases where different names and codes for the same company were firmly established, multiple codes may occur (e.g. RRI and SPT).

- In cases where an oil company has an own laboratory, the same code can be used for the laboratory and the company (*Sender*), if necessary with an extension to denote one particular of many company laboratories.
- If a company is not included in the list, the full official name shall be used.

3.2.5 Legal country codes

Legal values of *Country* (an attribute of *Sample Site*) are listed in the “LookupValue” sheet of the workbook. These are the Alpha-2 country codes (2 characters) according to ISO 3166.

Remarks

- Note that the code for United Kingdom is GB (not UK) and the code for Germany is DE (not GE, which is the code for Georgia).

3.2.6 Legal sample site names

Sample site names apply to *SiteName* (attribute of *Sample Site*). The dictionary table for sample site currently contains the Norwegian exploration and development wells and the name UNKNOWN.

Remarks

- The correct sample site name of a Norwegian well is the complete wellbore reference as defined by the NPD, but using the compact format without leading or embedded blanks or zeroes, except for one space between the well reference (items 1-4) and the remaining items (5-7). Both hyphens must be included. Examples: 1/3-8, 6407/6-3, 1/5-3 S, 2/4-14 R2, 15/9-19 SR2, 1/9-A-1 AH, 31/2-K-14 AY1H. An updated well list for the Norwegian continental shelf is available at NPD's web site[8].
- For sample sites other than wells, official names (if available) should be used, and the geographic coordinates should be reported (see Attribute list for Sample Site).
- For reference samples the site name UNKNOWN should be used and the "name" of the reference sample (e.g. NGS NSO-1) should be reported as OpRefNumber.

3.2.7 Legal site type codes

Legal values of *SiteType* (an attribute of *Sample Site*) are listed in the “LookupValue” sheet of the workbook.

Remarks

- The site type UNKNOWN should be used for the NGS[9] reference samples.

8 Fact-Pages > Downloads.

9 Norwegian Geochemical Standard

3.2.8 Legal sample type codes

Legal values of *SampleType* (an attribute of *Sample*) are listed in the "LookupValue" sheet of the workbook.

Remarks

- The operator company who dispatches a previously prepared fraction to a service company for analysis, is required to provide sample and fraction identification data in order to ensure correct reporting of the results (see general reporting rules in NIGOGA).
- The following sample types should be used for the NGS reference samples: OIL for NSO-1, OC for SR-1 and JR-1.

3.2.9 Legal lithology descriptors

Legal descriptors to be used in *SLithology* (an attribute of *Sample*) and *FLithology* (an attribute of *Fraction*) are listed in the "Lithology" sheet of the workbook. They include main lithologies (e.g. SLST), lithology modifiers/qualifiers (e.g. SLTY), colours (e.g. GY), colour shades (e.g. DK) and descriptors for grain size, grain shape, bedding and consistence (e.g. LAM). For syntax see below. The code numbers refer to the lithology attribute types listed in the "LithologyAttrType" sheet.

Remarks

- The purpose of the lithology characterisation is <u>not</u> a complete and detailed description, but a <u>searchable</u>, and therefore standardised, brief characterisation of the main constituents, the geochemically relevant accessories (e.g. bitumen, coal) and the visible additives. Should the type of mud (water-based, oil-based etc.) be known, this can be specified in a comment field.
- Suggested syntax:
  Main or major lithologies are separated by a slash (/). Each major lithology can be followed by a qualifier which is enclosed in parentheses. Each qualifier can contain several codes which must be separated by commas, except for colour shades if followed by a colour (e.g. LT GY). Qualifiers should be written in the following order: lithology, grain size, colour, other qualifier types).

  Examples:

A rock consisting of claystone and siltstone laminae	CLYST / SLST (LAM)
A claystone having a significant silt content (silty claystone)	CLYST (SLTY)
A sandstone with siliceous cement	SST (SIL, CMT)
Light grey and dark grey claystone	CLYST (LT GY) / CLYST (DK GY)
Red, coarse-grained sandstone + light grey, fine-grained sandstone + unspecified mud additive	SST (CGR, RD) / SST (FGR, LT GY) / ADD

- The lithology CLYST should be used for the NGS reference samples SR-1 and JR-1.

3.2.10 Legal fraction type codes

Legal values of *FractionType* (an attribute of *Fraction*) are listed in the "LookupValue" sheet of the workbook.

Remarks

- *FractionType* is a mandatory attribute and must always be defined even if the whole sample was analysed without preceding separation. Use BULK to denote the whole, untreated sample of any kind; use SDUN to denote crushed, unextracted sediment (e.g. for Rock-Eval analyses); use SDEX for crushed, pre-extracted sediment.

3.2.11 Legal analysis type codes

Legal values of *AnalType* (an attribute of *Analysis*) are listed in the "AnalysisType" sheet of the workbook.

Remarks

- In version 2.0 the differentiation between "raw" and "derived" data was abandoned as the parameter names are unique anyway. One new analysis type (ELEMENT) was introduced.

3.2.12 Legal analysis description codes

Legal values of *ADescription* (an attribute of *Analysis*) are listed in the "LookupValue" sheet of the workbook.

Remarks

- Note that the NIGOGA requires the reporting of results from control analyses of reference samples.

3.2.13 Legal detector codes

Legal values of *Detector* (an attribute of *Parameter*) are listed in the "LookupValue" sheet of the workbook.

Remarks

- *Detector* is mandatory for all types of analyses based on gas chromatographic separation of components (*AnalType* = GC).
- When using GC-IRMS, the detector code must refer to the isotope ratio determined.
- When using GC-MS, the detector code must be equal to the mass (m/z) of the detected ion or fragment (integer values). Lists of all ions to be reported are found in the NIGOGA.

3.2.14 Legal peak property codes

Legal values of *PeakProperty* (an attribute of *Parameter*, new in version 2.0) are listed in the "LookupValue" sheet of the workbook.

Remarks

- The new attribute *PeakProperty* replaces the formerly used suffixes to the parameter names (_H, _A, _N and _C).
- *PeakProperty* is mandatory for all types of analyses based on gas chromatographic separation of components (*AnalType* = GC).
- The *PeakProperty* value *ISOTRATIO* is introduced in version 2.0 to cater for isotopic analyses.

3.2.15 Legal parameter names

Legal values of *ParamName* (an attribute of *Parameter*) are listed in the “ParameterName” sheet of the workbook.

Remarks

- The list contains descriptions of the (geochemical) parameters that can be included in a GC-NPD-95 transfer file. This includes all parameters that have to be reported according to the NIGOGA, and additional parameters that may be useful.
- It also contains the units of measure which are implicit from the *ParamName* value. These are defined in the NIGOGA document (available at NPD's website) and are copied to the dictionary only for convenience of use. In case of a conflict between NIGOGA and the dictionary, NIGOGA overrules the dictionary.

4. PHYSICAL DATA FORMAT

4.1 Introduction

This chapter describes the ASCII representation of the GC-NPD-95 Data Transfer Format.

4.2 Transfer medium

All "usual" transfer media (diskette, CD-ROM, Video-8 tape etc.) and online transfer are acceptable. NPD has the right to refuse "exotic" or outdated transfer media.

The transfer medium must be properly labelled and the label must include all the mandatory attributes of the "Data Transfer Job" entity (see the sheet “Attributes” of the dictionary or Figure 6 of this document). Retransmission of (corrected) data sets must be clearly marked as such (using the mandatory attribute *DTJ_Status*).

4.3 General concepts and definitions

4.3.1 File type

GC-NPD-95 files are text files ("ASCII" or "flat" files). Word processor or spreadsheet file formats are not accepted.

As a rule, the transfer files should be uncompressed. If compression is required for specific reasons, sender and receiver should agree on the compression format (e.g. ZIP).

4.3.2 File name

The filename for reporting of data to NPD shall be GCH_RAW.ASC as described in the “Regulations relating to resource management in the petroleum activities, 18 June 2001, Section 24. Final reporting of geological and reservoir technical well data”[10].

4.3.3 Character set

The character set to be used is the 8-bit single-byte coded graphic character set defined in ISO 8859-1 (1998)[11]. This standard defines the following representation code for the Norwegian umlauts: Æ = 198, Ø = 216, Å = 197, æ = 230, ø = 248, å = 229. The degrees sign (°) is represented by the code 176.

4.3.4 List delimiters, text encapsulation and record length

GC-NPD-95 Version 2.0 allows exclusively character-delimited data formats, in which the values are separated by a character which is defined as a delimiter in the file header (file definition block, see Section 4.5.1 on File definition lines). The three legal delimiters are TAB (ANSI code 9), comma (ANSI code 44) and semicolon (ANSI code 59). Text strings containing any legal delimiter or double quotes (") must be encapsulated in double quotes

---

[10] For Internet addresses see the footnotes in the Preface.

[11] ISO/IEC 8859-1:1998 Information technology -- 8-bit single-byte coded graphic character sets -- Part 1: Latin alphabet No. 1 (available in English only). For updated version see www.iso.ch.

(ANSI code 34). Embedded double quotes, i.e. double quotes within a text string, should be avoided, but if they occur they must be duplicated (Figure 8).

The record length is theoretically unlimited, but for the sake of editing and printing the files for quality control the record length should not exceed the capacity of normal text editors and screens (without wrapping).

4.3.5 Decimal symbol and date format

The only legal decimal symbol is the dot (ANSI code 46).
The only legal date format is DD.MM.YYYY.

4.3.6 Case sensitivity

All keywords are case-sensitive[12].

4.3.7 Comment lines and blank records

The file can contain an unlimited number of *comment lines*. These are indicated by a double slash (//) in columns 1 and 2. Comment lines can be inserted at any point in the file. Characters following the double slash will be ignored by the reading program.

Blank records, i.e. records that contain only line feed (ANSI code 10) or carriage return followed by line feed (ANSI codes 13 and 10), can be inserted to increase readability.

```
-25.78,27.05.2001,Text code,25/4-6, Nothing in this line needs to be encapsulated in double quotes.
This text contains a full stop but no embedded delimiters or double quotes. No encapsulation required.
"This text contains a legal delimiter, hence it must be encapsulated in double quotes."
"This text contains another legal delimiter; hence it must be encapsulated in double quotes."
"This text contains a TAB character       (legal delimiter) and must be encapsulated in double quotes."
This text contains embedded ""double quotes"". These must be duplicated.
This text contains embedded 'single quotes'. These need no special treatment.
// This comment starts with a double slash and requires therefore no encapsulation.
-----
// The preceding line contains a block terminator ('-----').
```

Figure 8 Encapsulation of text. All records shown are correctly encapsulated.

4.3.8 Missing values

Missing numeric or text values shall be represented as either nothing (e.g. 25.3,,44.1,,,32.7) or space (e.g. 25.3, ,44.1, , , 32.7). Multiple spaces will be interpreted as one space.

4.4 Format structure

A GC-NPD-95 data transfer file contains data for exactly one Data Transfer Job. It consists of *data blocks*, i.e. identifiable groups of data within the physical data format (Figure 9). Each data block corresponds to exactly one entity type of the logical data model (cf. Chapter 2) and

[12] Case sensitivity is required in view of the possible future use of the (case-sensitive) XML format for data transfer.

is terminated by a *block terminator*, which consists of five consecutive hyphens in columns 1 to 5 (-----).
There exist three types of data blocks (*block types*), each containing characteristic elements (Figure 10):

(1) The ***file definition block***. It consists of *file definition lines*, which describe the attributes of the Data Transfer Job.

(2) The ***record definition block***. It consists of a *block definition line* followed by *record definition lines*, which describe the contents and structures of (a) the *value records* to follow this block, or (b) of the following record definition lines within the same record definition block (see Figure 19). Record definition blocks may define up to nine different *value assignment record types*, each containing values for a different set of parameters.

(3) The ***value assignment block***. It consists of *value assignment records*, which contain the actual geochemical values in the format(s) described in the last record definition block before the assignment block.

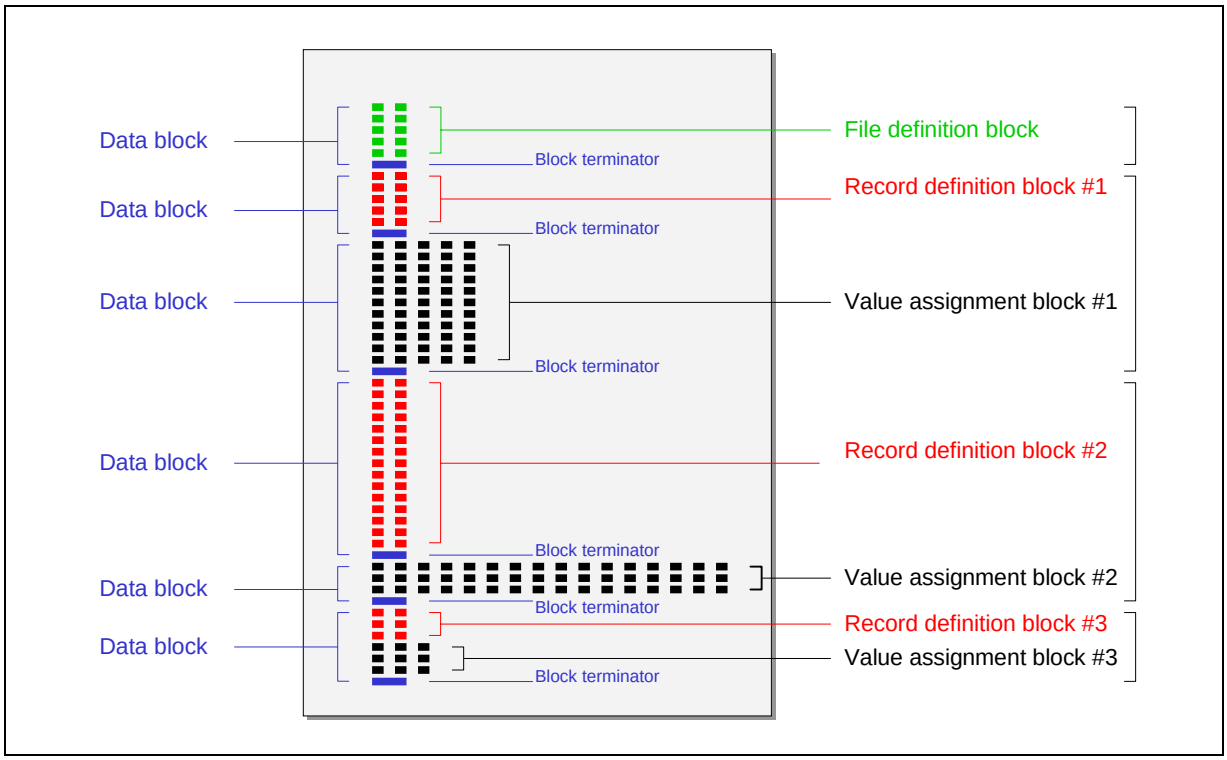


Figure 9 Structure of a GC-NPD-95 file.

The three block types must be arranged in the file according to the following rules:

1. A GC-NPD-95 file must contain exactly one file definition block, which must always be the first block in the file.
2. The file definition block is followed by up to 99 couples of record definition and value assignment blocks.

3. Each record definition block has exactly one corresponding value assignment block, and vice versa. Each record definition block must be directly followed by the corresponding value assignment block.
4. Each record definition block represents exactly one entity type in the logical data model (yellow boxes in Figure 6).
5. The block couples must be arranged according to the hierarchy of the entity types, i.e. parents must be defined before the children (e.g. Sample Site before Sample, Sample before Fraction and so on).

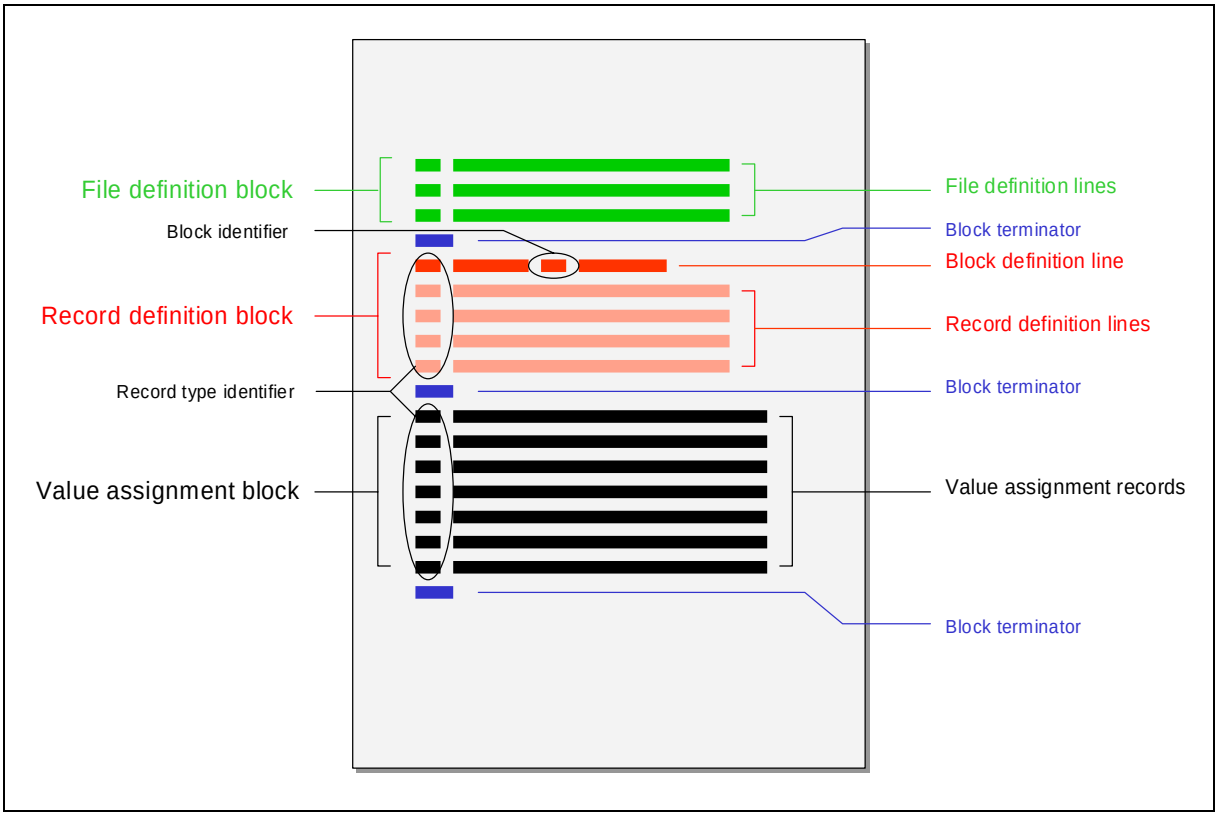


Figure 10 Elements of a File definition block, a Record definition block and a Value assignment block.

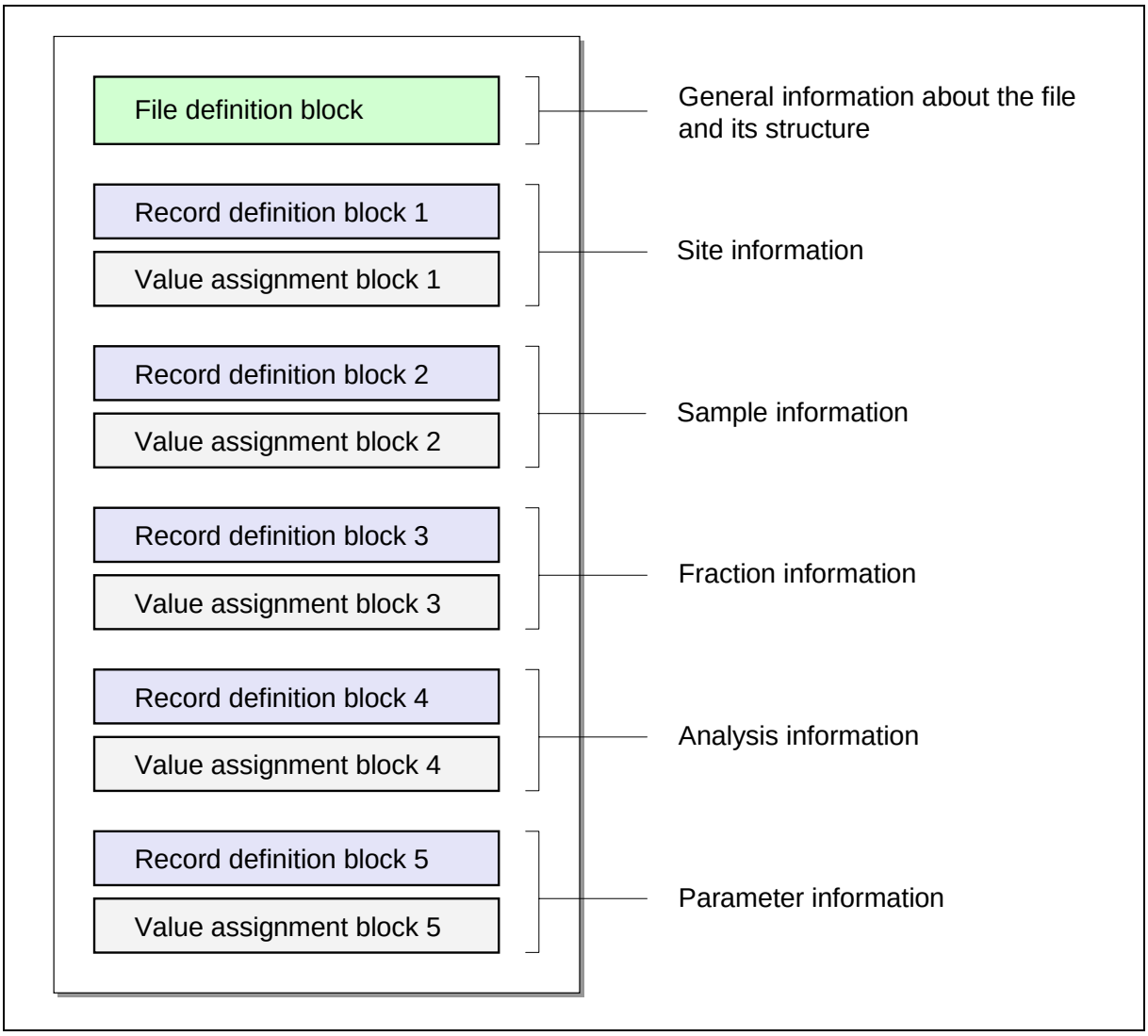


Figure 11 Arrangement of data blocks in a GC-NPD-95 file.

4.5 Record formats

Only the file definition lines and the block definition lines have predefined formats. The formats of all other records are defined in the transfer file itself. The *type* of each record is recognised by (a) the position in the file (applies only to the file definition block) or (b) by the *record-type identifier* which is indicated by two non-blank characters in columns 1 and 2 of each record.

4.5.1 File definition lines

File definition lines are recognised implicitly by their position between the top of file and the first block terminator. They are always in fixed, column-bound format. Each file definition line consists of an attribute name (columns 1-15) and an attribute value (column 16 to end of line; Figure 12). It is strongly recommended to place the format, version and dictionary definition lines at the top of the block, as shown in Figure 13).

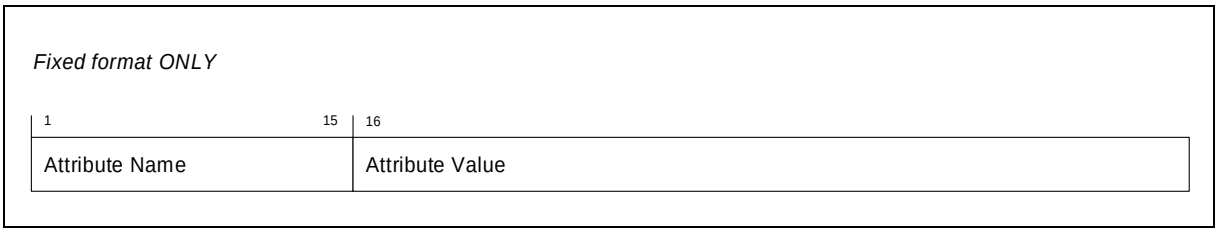


Figure 12 General format of a file definition line.

```
//..:....1....:....2....:....3....:
Format          GC-NPD-95
Version         2.0
Dictionary      GC-DIC-V2
Delimiter       COMMA
TransferID      OD_TEST-JUN95
TransferDate    13.06.1995
Sender          OILCOMPANY X
Contact         Kari Normann
Recipient       OD
```

Figure 13 Example of a file definition block. The delimiter must always be specified (red). The ruler (comment line, blue) is only for orientation and does not belong to the file definition block.

4.5.2 Block definition lines

The block definition line is the first line in a record definition block. It is recognised by the record-type identifier 00 (exactly two zeroes), which must be placed in the first two character positions of the record. The record-type identifier must be followed by the text string DEFINE BLOCK and a *block identifier* which consists of a two-digit number between 01 and 99 (leading zero to be included). A comment that describes the block content may be appended at the end of the block definition line. This comment is optional. Figure 14 shows the correct placement of these elements.

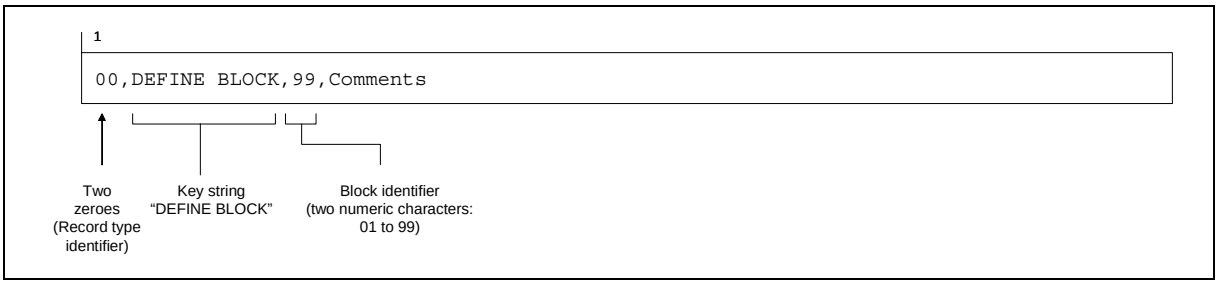


Figure 14 General format of a block definition line (comma-delimited format).

4.5.3 Record definition lines

Record definition lines are recognised by their position between a block definition line and the next block terminator. Record definition lines start with the record-type identifier L*n* where *n* designates a numeric character between 0 and 9.

Lines having identifiers between L1 and L9 (*format identifiers*) describe the contents (i.e. the parameter names) and formats of the corresponding *value assignment records* in the following value assignment block. It is therefore possible to define up to nine different record formats within the same data block, each having its own parameter set. Lines having the identifier L0 (zero) describe the content and format of *record definition lines* that immediately follow the L0 line within the present record definition block (see explanation below).

The normal format of a record definition line is described in Figure 15 and illustrated by the example in Figure 16. Note that the number *n* of the format identifier must increase from line to line within a record definition block. Gaps in the numbering are allowed but should generally be avoided.

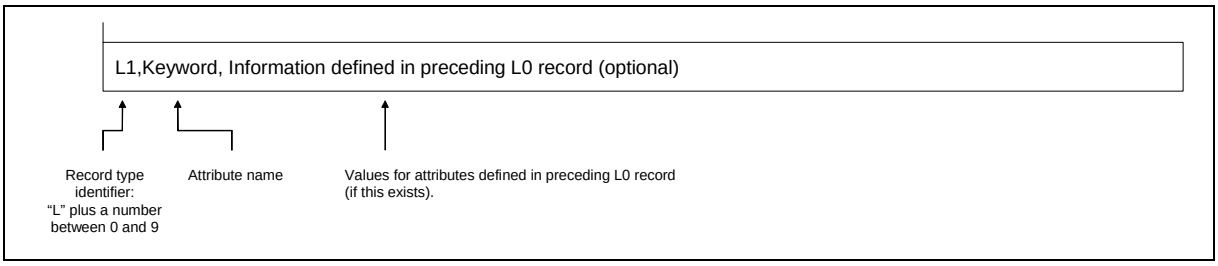


Figure 15 General format of a record definition line (comma-delimited format).

As up to nine different value assignment record formats (corresponding to the format identifiers L1 to L9) can be defined within a single record definition block, the data can be grouped and formatted in a flexible way.

```
00,DEFINE BLOCK,01,Sample Site (Well/Outcrop) data
L1,TransferID
L1,SiteID
L1,Country
L1,SiteType
L1,SiteName
L1,DatumName
L1,DatumElevation
L1,SSComments
-----
```

Figure 16 A record definition block that defines one value assignment record format.

```
// ------------------------------------------------------
// Record definition block, two record formats
// ------------------------------------------------------
00,DEFINE BLOCK,04,Sample
L1,SampleID
L1,UDepth
L1,LDepth
L1,SampleType
L1,LabRefNumber
L1,SCompany
L2,SLithology
L3,SComments
-----
// ------------------------------------------------------
// The corresponding value assigment records look as follows:
// ------------------------------------------------------
04,3512-0052,1230.00,1235.00,DC,H1235,Super Well Services
L2,CLYST / SLST (LAM)
L3,
04,3512-0052,1433.50,1433.70,COCH,H1236,Super Core Services
L2,SST (CALC)
L3,"Sample contaminated by oil-based mud, requires extraction before analysis"
-----
```

Figure 17 *A record definition block defining three different record formats in the same block (L1 = brown, L2 = pink, L3 = green).*

The names of the geochemical parameters (*ParamName*), the Detector type names (*Detector*) and the *PeakProperty* (i.e. area, height etc.) are, according to the logical data model in Figure 6, *attribute values* and would therefore normally be included in the value assignment block (Figure 18).

```
// ------------------------------------------------------------------
// Record definition block, parameter names in value assignment block
// ------------------------------------------------------------------
00,DEFINE BLOCK,16,Rock-Eval
L1,AnalID
L1,ParamName
L1,ParamValue
L1,ParamName
L1,ParamValue
L1,ParamName
L1,ParamValue
L1,ParamName
L1,ParamValue
-----
// ------------------------------------------------------
// Corresponding value assigment block
// ------------------------------------------------------
16,RE3512-0052-01,TOC,12.4,S1,0.35,S2,4.29,Tmax,438
16,RE3512-0103-01,TOC,3.86,S1,0.12,S2,2.63,Tmax,442
16,RE3512-0164-02,TOC,2.09,S1,0.08,S2,0.54,Tmax,436
-----
```

Figure 18 *Value assignment block containing parameter names.*

This allows almost unlimited freedom in the arrangement of the results, but may lead to unnecessary repetition of keywords in the value assignment blocks. To avoid such redundancy, GC-NPD-95 also allows the *ParamName*s to be declared in the record definition block. This can be done in two different ways:

Method 1 (explicit; Figure 19):
One of the record definition lines is assigned the *record-type identifier* L0, followed by the attribute name *ParamName*. This implies that the parameter names (i.e. the values of the attribute *ParamName*) will be included after the *ParamValue* keyword in the following record definition lines (until the end of the block). This method was defined in Version 1.0 of GC-NPD-95 but is discouraged because of its complexity.

```
//..:....1....:....2....:....3....:....4....:
00,DEFINE BLOCK,04,Rock-Eval
L1,AnalID
L0,ParamName
L1,ParamValue,S0
L1,ParamValue,S1
L1,ParamValue,S2
L1,ParamValue,TOC
L1,ParamValue,TMAX
-----
```

Figure 19 *A record definition block that declares parameter names using the explicit method. The third line (having the record-type identifier L0) defines that the following record definition lines will contain the values (S0, S1 etc.) of the* ParamName *attribute.*

Method 2 (implicit; Figure 20):
This method allows to omit the *ParamValue* keyword and to place the *ParamName*s (i.e. attribute values) directly in the record definition lines. This method is new in version 2 of GC-NPD-95 and applies only to the attributes *ParamName*, *Detector* and *PeakProperty* (see examples at the end of this chapter). It is recommended to use this "implicit" method rather than Method 1.

```
//..:....1....:....2....:....3....
00,DEFINE BLOCK,04,Rock-Eval
L1,AnalID
L1,S0
L1,S1
L1,S2
L1,TOC
L1,TMAX
-----
```

Figure 20 *A record definition block with declaration of parameter names using the implicit method. New in version 2.*

The *ParamName*s belonging to the same value assignment record can also be written consecutively, as shown in Figure 21 (simplified implicit method).

```
//..:....1....:....2....:....3....:....4....:
00,DEFINE BLOCK,04,Rock-Eval
L1,AnalID,S0,S1,S2,TOC,TMAX
-----
```

Figure 21 *A record definition block using the simplified implicit method. New in version 2.*

The value assignment block that corresponds to the definition blocks in Figure 19, Figure 20 and Figure 21 is shown in Figure 26.

Special rules for GC- and GC-MS data

GC- and GC-MS data require the specification of two attributes in addition to the analysis identifier: *Detector* and *PeakProperty*. *Detector* specifies the detector type (GC) or the detected ion mass (GC-MS) or isotope ratio (GC-IRMS). *PeakProperty* (i.e. height, area, concentration etc.) replaces the suffixes to the parameter names used in Version 1 (i.e. _H, _A, -C, _N) and thus drastically reduces the number of parameter names. The values of these two attributes must be specified before the first reported compound, and they apply to all following compounds until the end of the record (i.e. to only one level) or until new values are specified. These re-specifications must always include *Detector* and *PeakProperty* (Figure 22).

```
00,DEFINE BLOCK,32,GC-MS of saturated hydrocarbon fraction
L1,AnalID
L1,Detector          ← Detector (m/z 191; must be specified before the first ParamName)
L1,PeakProperty      ← PeakProperty (must be specified before the first ParamName)
L1,27Tm
L1,27Ts
L1,30ab
L1,Detector          ← new Detector (m/z 217)
L1,PeakProperty      ← PeakProperty (must be specified each time Detector is re-specified and vice versa)
L1,27dbS
L1,27dbR
L1,Detector          ← same Detector (m/z 217)
L1,PeakProperty      ← but new PeakProperty (peak area instead of peak height)
L1,27dbS
L1,27dbR
L2,Detector          ← new record format (L2), therefore Detector and PeakProperty must be re-specified,
L2,PeakProperty      ← although neither of them has changed
L2,29aaS
L2,29bbR
L2,29bbS
L2,29aaR
...
```

Figure 22 *The use of* Detector *and* PeakProperty. *Names of mandatory attributes in pink, parameter names in black. The value assignment record corresponding to this definition is shown in Figure 27.*

```
//..:....1....:....2....:....3....:....4....:....5....:....6....:....7....:....8...
.:....9....:...10....:...11....:...12....:...13....:...14
00,DEFINE BLOCK,32,GC-MS of saturated hydrocarbon fraction
L1,AnalID,Detector,PeakProperty,27Tm,27Ts,30ab,Detector,PeakProperty,27dbS,27dbR,De
tector,PeakProperty,27dbS,27dbR
L2,Detector,PeakProperty,29aaS,29bbR,29bbS,29aaR...
```

Figure 23 *The use of* Detector *and* PeakProperty *in alternative format. Names of mandatory attributes in pink, parameter names in black. Note that long records were wrapped by the word processor to fit into the frame. Corresponding value assignment record shown in Figure 27.*

Special rules for vitrinite reflectance data

Aggregated vitrinite reflectance data (mean reflectance, standard deviation etc.) require the specification of one attribute: *PopnNum*, which is the running number of the vitrinite population whose properties (mean reflectance, standard deviation, number of points measured) are reported. This attribute replaces the former parameter POPNNO which did not fully comply with the logical data model. The value of the *PopnNum* must be specified before the properties of the respective population.

As the number of measured vitrinite populations typically changes from sample to sample, it is most sensible and economic to report each population in a separate line (Figure 24, Figure 25).

```
00,DEFINE BLOCK,28,Vitrinite reflectance
L1,AnalID
L1,PopnNum      ← Population number must be specified before the first ParamName
L1,PopnMean
L1,PopnStdev
L1,PopnRead
L1,VRReliability
L1,VRQuality
...
-----
```

Figure 24 *The use of* PopnNum*. The value assignment block corresponding to this definition is shown in* Figure 28.

```
00,DEFINE BLOCK,28,Vitrinite reflectance
L1,AnalID,PopnNum,PopnMean,PopnStdev,PopnRead,VRReliability,VRQuality,...
-----
```

Figure 25 *The use of* PopnNum *in alternative format. Corresponding value assignment block shown in Figure 28.*

Reflectance values of individual vitrinite particles are reported without a population number. For format requirements see the data dictionary.

4.5.4 Value assignment records

Value assignment records are placed directly after the corresponding record definition block. Their record-type identifier depends on the format identifier (L1-L9) of the corresponding record definition line. Value assignment records of format L1 start with the *block identifier* which serves as a pointer to the corresponding record definition block and implicitly marks the record format as L1. Records of the formats L2 to L9 (if defined) start with the *format identifier*. These records must follow immediately after the first record (see Figure 17 and Figure 27).

```
//..:....1....:....2....:....3....
04,20004,0.15,1.04,0.12,0.35,445
```

Figure 26 *Example of a value assignment record. The ruler (comment line) is only for orientation and is not part of the value record. Given the definition shown in Figure 19 to Figure 21, this record belongs to Block number 04, the unique identification code of the analysis (AnalID, printed in red) is 20004, and values are reported for S0 (=0.15), S1 (=1.04), S2 (=0.12), TOC (=0.35), and TMAX (=445). The units of measure are defined in the NIGOGA and listed in the data dictionary table for ParamName.*

```
//..:....1....:....2....:....3....:....4....:....5....:....6....:....7....:....8....:
....9....:...10....:...11....:...12....:...13....:...14
32,2004A9280,191,HEIGHT,225.39,821.35,3276.29,217,HEIGHT,493.46,357.16,217,AREA,5724.
96,3967.64
L2,217,AREA,2634.87,3274.78,3421.57,2898.24 ...
```

Figure 27 *Value assignment record corresponding to the definition given in Figure 22 and Figure 23. Attribute values in pink, parameter values in black, value assignment record identifiers bolded. The ruler (comment line) is only for orientation and is not part of the value record. Note that long records were wrapped by the word processor to fit into the frame.*

```
//..:....1....:....2....:....3....:....4....:....5....:....6....:....7....:....8....:
28,2004A4536,1,0.63,0.05,21,good,main population
28,2004A4536,2,0.32,0.12,17,moderate,downfall?
28,2004A4536,3,1.20,0.28,5,poor,reworked
28,2004A6694,1,0.75,0.08,5,moderate,main population
28,2004A6694,2,0.28,0.10,3,poor,stained vitrinite
...
```

Figure 28 *Value assignment records corresponding to the definition given in Figure 24 and Figure 25. Vitrinite reflectance analysis of two samples having three and two populations, respectively. Attribute values in pink, parameter values in black. The ruler (comment line) is only for orientation and is not part of the value record.*

5. MAJOR CHANGES SINCE VERSION 1.0

This chapter gives a quick overview of the most important differences between version 1 and version 2 of the GC-NPD-95 format.

5.1 Specification document

- The document was completely revised, and a general introduction was added.
- The document is now available in PDF format at NPD's web site.
- Attribute lists and lists of legal attribute values were moved into an Excel® Workbook.
- Examples were substantially extended and moved to separate documents/files.

5.2 Logical data model

- Some changes were made in the data structure to simplify the physical format.
- A few new attributes were added, and some illogical names of entities, attributes and parameters have been changed to obtain consistency (see the lists below).
- The suffixes "(RAW)" and "(DERIVED)" were removed from the names of the analysis types.
- *Detector* and *POPNNO* (now *PopnNum*) are no longer parameters, but attributes of the *Parameter* entity type.
- *PeakProperty*, a new attribute of *Parameter*, was introduced to indicate which property of a GC peak is measured (peak height, peak area, etc.). This attribute replaces the formerly used suffixes _H, _A etc. of the parameter names for GC peaks and thus substantially reduces the number of peak parameter names.
- Isotope analysis of individual, GC-separated compounds is now consequently handled as a "GC" analysis, with the isotope ratio code as *Detector* type, the code "ISOTRATIO" as *PeakProperty* and the compound name as *ParameterName*.
- The AnalysisType "ELEMENT" was added.

- New attributes:

V2.0	Attribute of	Comment
Delimiter	Data Transfer Job	Character used as delimiter.between data values.
DTJComments	Data Transfer Job	Comment on the transfer job.
SLithologyFull	Sample	Optional field for full description in user format.
FLithologyFull	Fraction	Optional field for full description in user format.
PeakProperty	Parameter	Introduced to avoid the necessity of separate sets of parameter names (with different suffixes) for peak heights, peak areas etc.
Detector	Parameter	Changed from being a parameter to being an attribute of the *Parameter* entity.
PopnNum	Parameter	Previously called POPNNO; changed from being a parameter to being an attribute of the *Parameter* entity.

- Changed attribute or entity type names:

V2.0	V1.0	Comment
Sender	Laboratory	Attribute of Data Transfer Job. Changed as the sender of the data is not necessarily a laboratory.
Recipient	Client	Attribute of Data Transfer Job. Changed as the recipient of the data is not necessarily a “client”.
SiteType	Stype	Attribute of Sample Site. Stype was illogical as the prefix “S” otherwise is used for attributes of Sample.
SiteName	SName	Attribute of Sample Site. Same reason as for SType.
SampleType	SampType	Attribute of Sample. Changed to obtain consistency with SampleID and SampleDate.
AnalType	AnalName	Attribute of Analysis. Changed to obtain consistency with SiteType, SampleType and FractionType.
ADescription	Description	Attribute of Analysis. Changed to obtain consistency with other attributes starting with “A” and to avoid a generic term.
AMethod	Method	Attribute of Analysis. Same reason as for ADescription.
ADate	Date	Attribute of Analysis. Same reason as for ADescription.
Parameter	ParamValue	*ParamValue* is an attribute of *Parameter*. This change affects only Figure 6 and has no practical consequence for the transfer file.
PComments	COMMENT	*PComments* is now an attribute of *Parameter* and no longer itself a parameter. This allows to comment each individual result and makes the data structure more consistent.

5.3 Data dictionary

- The dictionary table for parameter names (earlier called "Parameter list") has been updated to meet the requirements of NIGOGA and to reflect the changes made in the logical data model. The list is now included in an Excel® workbook which also contains the attribute list.
- The dictionary table for lithology was revised, and a consistent syntax for lithology descriptions is suggested.
- Generic terms for unknown or poorly defined attribute values were added, and some definitions were sharpened.
- Spelling of some variables was changed to obtain consistency and lower case characters were introduced to allow easier reading of complex compound abbreviations.

5.4 Transfer media and physical format

- All usual transfer media and online transfer are now allowed.
- The column-bound format of V1.0 was replaced by a comma-, tab- or semicolon-delimited text format.

5.5 Validation program

- An Internet-based program for validation of GC-NPD-95 V2.0 files is provided for testing of files before submission. The program tests the logical structure and the physical format.