

INFORMATION SOCIETY TECHNOLOGIES  
(IST)  
PROGRAMME



## OpenMolGRID

### NTP DATA SPECIFICATION

---

Contract Reference:	<b>IST-2001-37238</b>
Document identifier:	<b>OpenMolGRID-1-D1.1d-0104-2-1-NTPDataSpec</b>
Date:	<b>07/01/2004</b>
Work package:	<b>WP1: Grid Data Warehousing of Molecular Structure – Property (Activity) Information</b>
Partner:	<b>UU</b>
Lead Partner:	<b>UU</b>
Document status:	<b>DRAFT</b>
Classification:	<b>PUBLIC</b>
Deliverable identifier:	<b>D1.1d</b>

---

**Abstract:** This document is to specify what transformations must take place to data before it is placed into the OpenMolGRID data warehouse MOLDW.

**Delivery Slip**

	<b>Name</b>	<b>Partner</b>	<b>Date</b>
<b>From</b>	Damian McCourt Jean Jing Werner Dubitzky	UU	15/09/03
<b>Verified by</b>	WPM	All	01/10/03
<b>Approved by</b>	G.H.F.Diercksen (TC) R.Ferenczi (QE)	OMC CGX	12/10/03 05/11/03

**Document Log**

<b>Issue</b>	<b>Date</b>	<b>Comment</b>	<b>Author</b>
0-0	28/08/03	First Version	Damian McCourt
0-1	05/09/03	Changed document template	Damian McCourt
1-0	10/09/03	Submitted to WPM	Damian McCourt Jean Jing Werner Dubitzky
2-0	15/09/03	Submitted for Authorisation	Damian McCourt Jean Jing Werner Dubitzky
2-0	05/11/03	Accepted for authorisation, but status to remain as draft	Damian McCourt
2-1	07/01/04	Updated due to the change of the document template (version 1.3)	Jean Jing

**Document Change Record**

<b>Issue</b>	<b>Item</b>	<b>Reason for Change</b>
0-1	Format Change	New deliverable template
2-0	Document updates	Document Authorised, but remains as draft
2-1	The format of the head file is changed	The standard template of the document is changed

**Files**

Files in this section relate to actual storage locations on the BSCW server located at <https://hermes.chem.ut.ee/bscw/bscw.cgi>. The URL below describes the location on BSCW from the root OpenMolGRID directory

<b>Software Products</b>	<b>User files / URL</b>
Word 2000/XP	OpenMolGRID/Workpackage 1/Deliverables/ OpenMolGRID-1-D1.1d-0104-2-1-NTPDataSpec

**Project information**

Project acronym:	OpenMolGRID
Project full title:	Open Computing GRID for Molecular Science and Engineering
Proposal/Contract no.:	IST-2001-37238
European Commission:	
Project Officer:	Annalisa BOGLIOLO
Address:	European Commission - DG Information Society F2 - Grids for Complex Problem Solving B-1049 Brussels Belgium
Office	BU31 4/79
Phone:	+32 2 295 8131
Fax:	+32 2 299 1749
E-mail	<a href="mailto:annalisa.bogliolo@cec.eu.int">annalisa.bogliolo@cec.eu.int</a>
Project Coordinator:	Mathilde ROMBERG
Address:	Forschungszentrum Jülich GmbH ZAM D-52425 Jülich Germany
Phone:	+49 2461 61 3703
Fax:	+49 2461 61 6656
E-mail	<a href="mailto:m.romberg@fz-juelich.de">m.romberg@fz-juelich.de</a>

## Contents

<b>1. INTRODUCTION</b> .....	<b>5</b>
1.1. PURPOSE AND SCOPE.....	5
1.2. OVERVIEW.....	5
1.3. DOCUMENT STRUCTURE.....	5
<b>2. IDENTIFICATION OF CHEMICALS</b> .....	<b>6</b>
<b>3. TOXICITY</b> .....	<b>8</b>
<b>4. CARCINOGENICITY</b> .....	<b>15</b>
<b>5. PHYSICO-CHEMICAL PROPERTIES(PCP)</b> .....	<b>17</b>
5.1. SPECIFIC GRAVITY.....	17
5.2. DENSITY.....	20
5.3. MELTING POINT .....	22
5.4. BOILING POINT.....	24
5.5. VAPOR PRESSURE .....	27
5.6. VAPOR DENSITY .....	29
5.7. DECOMPOSITION TEMPERATURE.....	31
5.8. SOLUBILITY .....	34
5.9. FLASH POINT .....	36
5.10. AUTOIGNITION TEMPERATURE.....	39
5.11. PH VALUE.....	41
5.12. VISCOSITY.....	44
5.13. REFRACTIVE INDEX.....	46
5.14. OCTANOL WATER PARTITION COEFFICIENT .....	48
5.15. PARTITION COEFFICIENT .....	51
<b>6. ADME RELATED PROPERTIES AND DESCRIPTORS</b> .....	<b>52</b>
<b>7. REFERENCES</b> .....	<b>55</b>

## **1. Introduction**

### **1.1. Purpose and Scope**

The purpose of this document is to specify what transformations must take place to data before it is placed into the OpenMolGRID data warehouse MOLDW. This document is intended to accompany deliverable D1.1a [1]. The information to be integrated into MOLDW is based on the information presented in deliverable D1.3 [2].

### **1.2. Overview**

Currently specifications for data transformations have not yet taken place. This document intends to initiate this process by providing a mapping of the data required as stated in deliverable D1.3 [2] to what is available in the NTP data source. A description of NTP is provided in document D1.4b [3]. Where possible data transformations have been stated, but other areas require further specification by other partners in the consortium.

### **1.3. Document Structure**

In addition to this section the document contains the following sections:

- Section 2 – identification of chemicals
- Section 3 – Toxicity
- Section 4 – Carcinogenicity
- Section 5 – Physico-Chemical Properties(PCP)
- Section 6 – ADME Related Properties and Descriptors

## 2. Identification of Chemicals

Name	Type	Description	Present In Source	Format	Conversion Details
CAS (unique identifier)	Integer	The Chemical Abstract Service (CAS) number. This is an integer without hyphens or spaces	Present or not (null) as data 'CAS NUMBER' under 'IDENTIFIERS'	xxxxxx-xx-x e.g. 100-00-5  Digits with two hyphens in the digit. The last is always one digit, the middle two digits, and the first is between two and six digits.	xxxxxxxxx e.g. 100005  Without the two hyphens. The number consists of five to nine digits.
Molecular Weight	Float	The Atomic Mass Unit (amu) associated with the CAS.	Present or not (null) as data 'MOLECULAR WEIGHT' under 'PHYSICAL CHEMICAL DATA'	157.56	No changes

Chemical Formula	String	The chemical formula associated with the CAS	Present or not (null) as data 'CHEMICAL FORMULA' under 'IDENTIFIERS'	String consists of letters and digits e.g. C6H4ClNO2	No changes
Chemical Names	List of Strings	The names by which the chemical is known	Present as data 'BASE CHEMICAL NAME', 'PRIMARY NAME', and 'SYNONYMS' under 'IDENTIFIERS'	String (array) Each element stores one chemical name.	String array stores the value of 'BASE CHEMICAL NAME', value of 'PRIMARY NAME', and values 'SYNONYMS' in list.

### 3. Toxicity

Name	Type	Description	Present In Source	Format	Conversion Details
CAS (unique identifier)	Integer	The Chemical Abstract Service (CAS) number. This is an integer without hyphens or spaces	Present or not (null) as data 'CAS NUMBER' under 'IDENTIFIERS'	xxxxxx-xx-x e.g. 100-00-5  Digits with two hyphens in the digit. The last is always one digit, the middle two digits, and the first is between two and six digits.	xxxxxxxx e.g. 100005  Without the two hyphens. The number consists of five to nine digits.



Target Names	Species	List of Strings	A list of names associated with the target species	Present or not (null) as data 'specie' under 'TOXICITY'	String array Each of element stores one of species by using there letters, e.g. brd=domestic_bird bwd=wild_bird_species cat=cat chd=child ckn=chicken ctl=cattle dog=dog dck=duck frg=frog gpg=guinea_pig grh=grasshopper hmn=human hor=horse/donkey mam=mammal man=man mky=monkey mus=mouse pgn=pigeon rat=rat rbt=rabbit wmn=woman	Need proper specification/mapping to species name for all sources
IST-2001-37238			<b>PUBLIC</b>	9 / 55		

End Point Type	String	The type of toxicity measure used in this protocol e.g. LC <sub>50</sub> , LD <sub>50</sub> .	Present or not (null) as data 'typ. dose' under 'TOXICITY'	Four character string e.g. LC50 LD50 LDLo LCLo	No change Need spec of endpoints to be considered
Dose Metric	Float	The dose of chemical studied (for aquatic toxicity the compound is in the tank, and the fish exposed) measured in milligrams per kilogram (mg/kg) or milligrams per litre (mg/l).	Present or not (null) as data 'amount' under 'TOXICITY'	Float e.g. 500	May need conversion dependant on units

Dose Metric Units	String	The units associated with the Metric dose	Present or not (null) as data 'units' under 'TOXICITY'	String with different unit (lower or upper cases mixture)	Need to be standardised.
				e.g. gm/kg gm/kg/4D mg/kg MG/KG mg/kg/4W-I mg/m3 mg/m3/4H ug/kg pph/5M ppm ppm/3D ppm/30M ppm/4H ppm/8H ppm/2M ppm/10s	mg/kg mg/l

Dose Mol	Float	The dose of chemical studied measured in millimols per kilogram (mmol/kg) or millimols per litre (mmol/l).	Not present		
Dose Mol Units	String	The units associated with the Mol dose	Not present		
Exposure Time	Float	The amount of time the target species was exposed to the chemical. This is measured in hours.	Not present in raw format but present in unit		Can be derives from units
Protocol Details	Text	General information associated with the protocol. There is no particular format associated with this.	Unknown at present		

Mode of Action USA	String	<p>The Mode of Action of the chemical. These are classes based on the EPA (Duluth, USA) standard. These are as follows:</p> <ul style="list-style-type: none"> <li>• Non Polar Narcosis (Base Line Narcosis)</li> <li>• Polar Narcosis (Narcosis II)</li> <li>• Narcosis III (Ester/ Acrylate compounds)</li> <li>• Oxidative Phosphorilation uncoupling</li> <li>• Respiratory inhibition</li> <li>• Electrophile and proelectrophile reactivity</li> <li>• Acetilcholinesterase inhibition</li> <li>• Central nervous system seizure responses</li> </ul>	Not present		
Mode of Action EU	String	<p>The mechanism of action of the chemical. These are classes based on the EU (Netherlands) standard. These classes are as follows:</p> <ul style="list-style-type: none"> <li>• Non Polar Narcosis</li> <li>• Polar Narcosis</li> <li>• Reactive</li> <li>• Receptor Mediated</li> </ul>	Not present		

Author(s)	List of Strings	The author who reported the toxicity measure.	Present or not (null) as under 'SOURCES' It/they appear in the text of SOURCES	Surname, Initial. e.g. Sax, N. I.	It would be difficult to extract Author from the text by using programs as it has no certain patterns.
Year	Integer	The year in which the author reported the toxicity measure.	Present or not (null) as under 'SOURCES' It/they appear in the text of SOURCES	YYYY e.g. 1969	It would be difficult to extract Year from the text by using programs as it has no certain patterns.
Database	String	The database from which the toxicity measure was obtained.		String	NTP

#### 4. Carcinogenicity

Name	Type	Description	Present In Source	Format	Conversion Details
CAS (unique identifier)	Integer	The Chemical Abstract Service (CAS) number. This is an integer without hyphens or spaces	Present or not (null) as data 'CAS NUMBER' under 'IDENTIFIERS'	xxxxxx-xx-x e.g. 100-00-5 Digits with two hyphens in the digit. The last is always one digit, the middle two digits, and the first is between two and six digits.	xxxxxxxxx e.g. 100005 Without the two hyphens. The number consists of five to nine digits.

Carcinogenicity	String	<p>Classification of carcinogenicity according to classes proposed by the World Health Organisation (International Agency for Research on Cancer- IARC). These classes are as follows:</p> <p>1 – Carcinogenic to human</p> <p>2a – Probably Carcinogenic to human</p> <p>2b – Possibly Carcinogenic to human</p> <p>3 – Unknown</p> <p>4 – Non-carcinogenic to human</p>	<p>Present or not (null) as data 'CARCINOGENICITY' subtitle 'Review' under 'TOXICITY'</p>	<p>String</p> <p>e.g.</p> <p>IARC Cancer Review: Animal Sufficient Evidence</p> <p>IARC Cancer Review: Human Inadequate Evidence</p> <p>IARC possible human carcinogen (Group 2B) [610]</p>	No change
-----------------	--------	---	---	---	-----------



## 5. Physico-Chemical Properties(PCP)

### 5.1. Specific Gravity

Name	Type	Description	Present In Source	Format	Conversion Details
CAS (unique identifier)	Integer	The Chemical Abstract Service (CAS) number. This is an integer without hyphens or spaces	Present or not (null) as data 'CAS NUMBER' under 'IDENTIFIERS'	xxxxxx-xx-x e.g. 100-00-5  Digits with two hyphens in the digit. The last is always one digit, the middle two digits, and the first is between two and six digits.	xxxxxxxxx e.g. 100005  Without the two hyphens. The number consists of five to nine digits.
PCP Name	String	Specific Gravity	Present as data 'SPECIFIC GRAVITY' under 'PHYSICAL CHEMICAL DATA'	string e.g. Specific Gravity	No changes

PCP Value	Float	Value measured for Specific Gravity	Present or not (null) as data 'SPECIFIC GRAVITY' under 'PHYSICAL CHEMICAL DATA'	The value together with the temperature and pressure measured, and the references e.g. 1.181 @ 20/4 C [026,043]	Extract the float value
PCP Unit	String	Unit of measured PCP	No Unit		
PCP Method	String	Method of measurement	Not present		Not present
PCP Temp	Float	Condition of the measurement	Present or not (null) as data 'SPECIFIC GRAVITY' under 'PHYSICAL CHEMICAL DATA'	e.g. 20/4	How to deal with range
PCP Temp Unit	String	Condition of the measurement, unit for temperature	Present or not (null) as data 'SPECIFIC GRAVITY' under 'PHYSICAL CHEMICAL DATA'	C or c	C
PCP Pressure	Float	Condition of the measurement, pressure	Not present		
PCP Pressure Unit	String	Condition of the measurement, unit for pressure	Not present		
PCP solvent	String	Condition of the measurement solvent	Not present		

PCP Concentration	Float	Condition of the measurement, concentration	Not present		
PCP Concentration Unit	String	Condition of the measurement, unit for concentration	Not present		
PCP Reference	Text	Literature source for respective PCP	Present or not (null) as data 'SPECIFIC GRAVITY' under 'PHYSICAL CHEMICAL DATA'	Reference number in bracket, separate by commas e.g. [017,058,205,275], which appear in the final under SOURCES	Extract from 'SOURCES'

## 5.2. Density

Name	Type	Description	Present In Source	Format	Conversion Details
CAS (unique identifier)	Integer	The Chemical Abstract Service (CAS) number. This is an integer without hyphens or spaces	Present or not (null) as data 'CAS NUMBER' under 'IDENTIFIERS'	xxxxxx-xx-x e.g. 100-00-5  Digits with two hyphens in the digit. The last is always one digit, the middle two digits, and the first is between two and six digits.	xxxxxxxxx e.g. 100005  Without the two hyphens. The number consists of five to nine digits.
PCP Name	String	DENSITY	Present as data 'DENSITY' under 'PHYSICAL CHEMICAL DATA'	string e.g. DENSITY	No changes
PCP Value	Float	Value measured for DENSITY	Present or not (null) as data 'DENSITY' under 'PHYSICAL CHEMICAL DATA'	Float e.g. 2.031	Dependant on standard unit.

PCP Unit	String	Unit of measured PCP	Present or not (null) as data 'DENSITY' under 'PHYSICAL CHEMICAL DATA'	g/mL g/cm <sup>3</sup>	
PCP Method	String	Method of measurement	Not present		Not present
PCP Temp	Float	Condition of the measurement	Present or not (null) as data 'DENSITY' under 'PHYSICAL CHEMICAL DATA'	Float number e.g. 18.5	No change
PCP Temp Unit	String	Condition of the measurement, unit for temperature	Present or not (null) as data 'DENSITY' under 'PHYSICAL CHEMICAL DATA'	C or c	C
PCP Pressure	Float	Condition of the measurement, pressure	Not present		
PCP Pressure Unit	String	Condition of the measurement, unit for pressure	Not present		
PCP solvent	String	Condition of the measurement solvent	Not present		
PCP Concentration	Float	Condition of the measurement, concentration	Not present		
PCP Concentration Unit	String	Condition of the measurement, unit for concentration	Not present		

PCP Reference	Text	Literature source for respective PCP	Present or not (null) as data 'DENSITY' under 'PHYSICAL CHEMICAL DATA'	Reference number in bracket e.g. [017,058,205,275], which appear in the final under SOURCES	Extract from 'SOURCES'
---------------	------	--------------------------------------	--	---	------------------------

**5.3. Melting Point**

Name	Type	Description	Present In Source	Format	Conversion Details
CAS (unique identifier)	Integer	The Chemical Abstract Service (CAS) number. This is an integer without hyphens or spaces	Present or not (null) as data 'CAS NUMBER' under 'IDENTIFIERS'	xxxxxx-xx-x e.g. 100-00-5  Digits with two hyphens in the digit. The last is always one digit, the middle two digits, and the first is between two and six digits.	xxxxxxxxx e.g. 100005  Without the two hyphens. The number consists of five to nine digits.

PCP Name	String	MP	Present as data 'MP' under 'PHYSICAL CHEMICAL DATA'	string e.g. MP	No changes
PCP Value	Float	Value measured for Melting Point	Present or not (null) as data 'MP' under 'PHYSICAL CHEMICAL DATA'	Float e.g. -25.6	No change
PCP Unit	String	Unit of measured PCP	Present or not (null) as data 'MP' under 'PHYSICAL CHEMICAL DATA'	C or c	C
PCP Method	String	Method of measurement	Not present		Not present
PCP Temp	Float	Condition of the measurement	Not Present		
PCP Temp Unit	String	Condition of the measurement, unit for temperature	Not Present		
PCP Pressure	Float	Condition of the measurement, pressure	Not present		
PCP Pressure Unit	String	Condition of the measurement, unit for pressure	Not present		
PCP solvent	String	Condition of the measurement solvent	Not present		
PCP Concentration	Float	Condition of the measurement, concentration	Not present		

PCP Concentration Unit	String	Condition of the measurement, unit for concentration	Not present		
PCP Reference	Text	Literature source for respective PCP	Present or not (null) as data 'MP' under 'PHYSICAL CHEMICAL DATA'	Reference number in bracket e.g. [017,058,205,275], which appear in the final under SOURCES	Extract from 'SOURCES'

**5.4. Boiling Point**



Name	Type	Description	Present In Source	Format	Conversion Details
CAS (unique identifier)	Integer	The Chemical Abstract Service (CAS) number. This is an integer without hyphens or spaces	Present or not (null) as data 'CAS NUMBER' under 'IDENTIFIERS'	xxxxxx-xx-x e.g. 100-00-5  Digits with two hyphens in the digit. The last is always one digit, the middle two digits, and the first is between two and six digits.	xxxxxxxxx e.g. 100005  Without the two hyphens. The number consists of five to nine digits.
PCP Name	String	BP	Present as data 'BP' under 'PHYSICAL CHEMICAL DATA'	string e.g. BP	No changes
PCP Value	Float	Value measured for BP	Present or not (null) as data 'BP' under 'PHYSICAL CHEMICAL DATA'	Float number e.g. 117.9	No change
PCP Unit	String	Unit of measured PCP	Present or not (null) as data 'BP' under 'PHYSICAL CHEMICAL DATA'	C or c	C

PCP Method	String	Method of measurement	Not present		Not present
PCP Temp	Float	Condition of the measurement	Not present		
PCP Temp Unit	String	Condition of the measurement, unit for temperature	Not present		
PCP Pressure	Float	Condition of the measurement, pressure	Present or not (null) as data 'BP' under 'PHYSICAL CHEMICAL DATA'	Float number e.g. 760	No change
PCP Pressure Unit	String	Condition of the measurement, unit for pressure	Present or not (null) as data 'BP' under 'PHYSICAL CHEMICAL DATA'	mm Hg mm/Hg mm hg	Inconsistency between representation of unit in files
PCP solvent	String	Condition of the measurement solvent	Not present		
PCP Concentration	Float	Condition of the measurement, concentration	Not present		
PCP Concentration Unit	String	Condition of the measurement, unit for concentration	Not present		

PCP Reference	Text	Literature source for respective PCP	Present or not (null) as data 'BP' under 'PHYSICAL CHEMICAL DATA'	Reference number in bracket e.g. [017,058,205,275], which appear in the final under SOURCES	Extract from 'SOURCES'
---------------	------	--------------------------------------	---	---	------------------------

**5.5. Vapor Pressure**

Name	Type	Description	Present In Source	Format	Conversion Details
CAS (unique identifier)	Integer	The Chemical Abstract Service (CAS) number. This is an integer without hyphens or spaces	Present or not (null) as data 'CAS NUMBER' under 'IDENTIFIERS'	xxxxxx-xx-x e.g. 100-00-5  Digits with two hyphens in the digit. The last is always one digit, the middle two digits, and the first is between two and six digits.	xxxxxxxxx e.g. 100005  Without the two hyphens. The number consists of five to nine digits.

PCP Name	String	Vapor pressure	Present under 'VOLATILITY'	string e.g. Vapor pressure	No changes
PCP Value	Float	Value measured for Vapor pressure	Present or not (null) as data 'Vapor pressure' under 'VOLATILITY'	Float number e.g. 10 mm Hg @ 16.6 C [038,043,051]; 40 mm Hg @ 42.0 C [038]	Extract the float numbers from text
PCP Unit	String	Unit of measured PCP	Present or not (null) as data 'Vapor pressure' under 'VOLATILITY'	mm Hg mm/Hg mm hg	Inconsistency between representation of unit in files
PCP Method	String	Method of measurement	Not present		Not present
PCP Temp	Float of array	Condition of the measurement	Present or not (null) as data 'Vapor pressure' under 'VOLATILITY'	10 mm Hg @ 16.6 C [038,043,051]; 40 mm Hg @ 42.0 C [038]	Extract the float numbers from text
PCP Temp Unit	String	Condition of the measurement, unit for temperature	Present or not (null) as data 'Vapor pressure' under 'VOLATILITY'	C or c	C
PCP Pressure	Float	Condition of the measurement, pressure	Not present		
PCP Pressure Unit	String	Condition of the measurement, unit for pressure	Not present		
PCP solvent	String	Condition of the measurement solvent	Not present		

PCP Concentration	Float	Condition of the measurement, concentration	Not present		
PCP Concentration Unit	String	Condition of the measurement, unit for concentration	Not present		
PCP Reference	Text	Literature source for respective PCP	Present or not (null) as data 'Vapor pressure' under 'VOLATILITY'	Reference number in bracket e.g. [017,058,205,275], which appear in the final under SOURCES	Extract from 'SOURCES'

**5.6. Vapor Density**

Name	Type	Description	Present In Source	Format	Conversion Details
CAS (unique identifier)	Integer	The Chemical Abstract Service (CAS) number. This is an integer without hyphens or spaces	Present or not (null) as data 'CAS NUMBER' under 'IDENTIFIERS'	xxxxxx-xx-x e.g. 100-00-5  Digits with two hyphens in the digit. The last is always one digit, the middle two digits, and the first is between two and six digits.	xxxxxxxxx e.g. 100005  Without the two hyphens. The number consists of five to nine digits.
PCP Name	String	Vapor density	Present under 'VOLATILITY'	string e.g. Vapor density	No change
PCP Value	Float	Value measured for Vapor density	Present or not (null) as data 'Vapor density' under 'VOLATILITY'	Float number e.g. 3.29	No change
PCP Unit	String	Unit of measured PCP	Present or not (null) as data 'Vapor density' under 'VOLATILITY'	No unit	
PCP Method	String	Method of measurement	Not present		Not present

PCP Temp	Float of array	Condition of the measurement	Not present		
PCP Temp Unit	String	Condition of the measurement, unit for temperature	Not present		
PCP Pressure	Float	Condition of the measurement, pressure	Not present		
PCP Pressure Unit	String	Condition of the measurement, unit for pressure	Not present		
PCP solvent	String	Condition of the measurement solvent	Not present		
PCP Concentration	Float	Condition of the measurement, concentration	Not present		
PCP Concentration Unit	String	Condition of the measurement, unit for concentration	Not present		
PCP Reference	Text	Literature source for respective PCP	Present or not (null) as data 'Vapor density' under 'VOLATILITY'	Reference number in bracket e.g. [017,058,205,275], which appear in the final under SOURCES	Extract from 'SOURCES'

**5.7. Decomposition Temperature**

Name	Type	Description	Present In Source	Format	Conversion Details
CAS (unique identifier)	Integer	The Chemical Abstract Service (CAS) number. This is an integer without hyphens or spaces	Present or not (null) as data 'CAS NUMBER' under 'IDENTIFIERS'	xxxxxx-xx-x e.g. 100-00-5  Digits with two hyphens in the digit. The last is always one digit, the middle two digits, and the first is between two and six digits.	xxxxxxxxx e.g. 100005  Without the two hyphens. The number consists of five to nine digits.
PCP Name	String	Decomposition Temperature	Present or not (null) as data 'BP' under 'PHYSICAL CHEMICAL DATA'	string e.g. Decomposes @ 840 C in the case 10101970.txt	Mapping the existing name
PCP Value	Float	Value measured for Decomposition Temperature	Present or not (null) as data 'BP' under 'PHYSICAL CHEMICAL DATA'	Float number e.g. 840	No change
PCP Unit	String	Unit of measured PCP	Present or not (null) as data 'BP' under 'PHYSICAL CHEMICAL DATA'	C or c	C



PCP Method	String	Method of measurement	Not present		Not present
PCP Temp	Float of array	Condition of the measurement	Not present		
PCP Temp Unit	String	Condition of the measurement, unit for temperature	Not present		
PCP Pressure	Float	Condition of the measurement, pressure	Not present		
PCP Pressure Unit	String	Condition of the measurement, unit for pressure	Not present		
PCP solvent	String	Condition of the measurement solvent	Not present		
PCP Concentration	Float	Condition of the measurement, concentration	Not present		
PCP Concentration Unit	String	Condition of the measurement, unit for concentration	Not present		
PCP Reference	Text	Literature source for respective PCP	Present or not (null) as data 'BP' under 'PHYSICAL CHEMICAL DATA'	Reference number in bracket e.g. [017,058,205,275], which appear in the final under SOURCES	Extract from 'SOURCES'

**5.8. Solubility**

Name	Type	Description	Present In Source	Format	Conversion Details
CAS (unique identifier)	Integer	The Chemical Abstract Service (CAS) number. This is an integer without hyphens or spaces	Present or not (null) as data 'CAS NUMBER' under 'IDENTIFIERS'	xxxxxx-xx-x e.g. 100-00-5  Digits with two hyphens in the digit. The last is always one digit, the middle two digits, and the first is between two and six digits.	xxxxxxxxx e.g. 100005  Without the two hyphens. The number consists of five to nine digits
PCP Name	String	SOLUBILITIES	Present or not (null) as data 'SOLUBILITIES' under 'PHYSICAL CHEMICAL DATA'	string e.g. SOLUBILITIES	No change
PCP Value	Float	Value measured for solubilities	Present or not (null) as data 'SOLUBILITIES' under 'PHYSICAL CHEMICAL DATA'	Range float number e.g. 1-5 or >=100	Inequalities will be stored in accordance with the description of the Property entity of the data model [1]

PCP Unit	String	Unit of measured PCP	Present or not (null) as data 'BP' under 'PHYSICAL CHEMICAL DATA'	C or c	C
PCP Method	String	Method of measurement	Not present		Not present
PCP Temp	Float	Condition of the measurement	Present or not (null) as data 'SOLUBILITIES' under 'PHYSICAL CHEMICAL DATA'	Float e.g. 22	No change
PCP Temp Unit	String	Condition of the measurement, unit for temperature	Present or not (null) as data 'SOLUBILITIES' under 'PHYSICAL CHEMICAL DATA'	C or c	C
PCP Pressure	Float	Condition of the measurement, pressure	Not present		
PCP Pressure Unit	String	Condition of the measurement, unit for pressure	Not present		
PCP solvent	String	Condition of the measurement solvent	Present or not (null) as data 'SOLUBILITIES' and 'OTHER SOLVENTS' under 'PHYSICAL CHEMICAL DATA'	e.g. WATER, ETHANOL	No change
PCP Concentration	Float	Condition of the measurement, concentration	Present or not (null) as data 'SOLUBILITIES' and 'OTHER SOLVENTS' under 'PHYSICAL CHEMICAL DATA'	Float e.g. 25	No change

PCP Concentration Unit	String	Condition of the measurement, unit for concentration	Present or not (null) as data 'SOLUBILITIES' and 'OTHER SOLVENTS' under 'PHYSICAL CHEMICAL DATA'	%	No change
PCP Reference	Text	Literature source for respective PCP	Present or not (null) as data 'SOLUBILITIES' under 'PHYSICAL CHEMICAL DATA'	Reference number in bracket e.g. [017,058,205,275], which appear in the final under SOURCES	Extract from 'SOURCES'

**5.9. Flash Point**

Name	Type	Description	Present In Source	Format	Conversion Details
CAS (unique identifier)	Integer	The Chemical Abstract Service (CAS) number. This is an integer without hyphens or spaces	Present or not (null) as data 'CAS NUMBER' under 'IDENTIFIERS'	xxxxxx-xx-x e.g. 100-00-5  Digits with two hyphens in the digit. The last is always one digit, the middle two digits, and the first is between two and six digits.	xxxxxxxxx e.g. 100005  Without the two hyphens. The number consists of five to nine digits.
PCP Name	String	FLASH POINT	Present or not (null) as data 'FLAMMABILITY (FLASH POINT)' under 'PHYSICAL CHEMICAL DATA'	string e.g. FLAMMABILITY(FLASH POINT)	FLASH POINT
PCP Value	Float	Value measured for Flash Point	Present or not (null) as data 'FLAMMABILITY(FLASH POINT)' under 'PHYSICAL CHEMICAL DATA'	Float number e.g. >93.3 C (>200 F)	Inequalities will be stored in accordance with the description of the Property entity of the data model [1]

PCP Unit	String	Unit of measured PCP	Present or not (null) as data 'FLAMMABILITY(FLASH POINT)' under 'PHYSICAL CHEMICAL DATA'	C or F	Change to C
PCP Method	String	Method of measurement	Not present		Not present
PCP Temp	Float	Condition of the measurement	Not present		
PCP Temp Unit	String	Condition of the measurement, unit for temperature	Not present		
PCP Pressure	Float	Condition of the measurement, pressure	Not present		
PCP Pressure Unit	String	Condition of the measurement, unit for pressure	Not present		
PCP solvent	String	Condition of the measurement solvent	Not present		
PCP Concentration	Float	Condition of the measurement, concentration	Not present		
PCP Concentration Unit	String	Condition of the measurement, unit for concentration	Not present		

PCP Reference	Text	Literature source for respective PCP	Present or not (null) as data 'FLAMMABILITY(FLASH POINT)' under 'PHYSICAL CHEMICAL DATA'	Reference number in bracket e.g. [017,058,205,275], which appear in the final under SOURCES	Extract 'SOURCES' from
---------------	------	--------------------------------------	--	---	------------------------

### 5.10. Autoignition Temperature

Name	Type	Description	Present In Source	Format	Conversion Details
CAS (unique identifier)	Integer	The Chemical Abstract Service (CAS) number. This is an integer without hyphens or spaces	Present or not (null) as data 'CAS NUMBER' under 'IDENTIFIERS'	xxxxxx-xx-x e.g. 100-00-5  Digits with two hyphens in the digit. The last is always one digit, the middle two digits, and the first is between two and six digits.	xxxxxxxxx e.g. 100005  Without the two hyphens. The number consists of five to nine digits.

PCP Name	String	AUTOIGNITION TEMPERATURE	Present or not (null) as data 'FLAMMABILITY (FLASH POINT)' under 'PHYSICAL CHEMICAL DATA'	Appears in the text of FLAMMABILITY. e.g. The autoignition temperature of this chemical	Change AUTOIGNITION TEMPERATURE to
PCP Value	Float	Value measured for Autoignition Temperature	Present or not (null) as data 'FLAMMABILITY(FLASH POINT)' under 'PHYSICAL CHEMICAL DATA'	Float number e.g. 231 C (448 F)	Extract the float that has the C unit
PCP Unit	String	Unit of measured PCP	Present or not (null) as data 'FLAMMABILITY(FLASH POINT)' under 'PHYSICAL CHEMICAL DATA'	C or F	Change to C
PCP Method	String	Method of measurement	Not present		Not present
PCP Temp	Float	Condition of the measurement	Not present		
PCP Temp Unit	String	Condition of the measurement, unit for temperature	Not present		
PCP Pressure	Float	Condition of the measurement, pressure	Not present		



PCP Pressure Unit	String	Condition of the measurement, unit for pressure	Not present		
PCP solvent	String	Condition of the measurement solvent	Not present		
PCP Concentration	Float	Condition of the measurement, concentration	Not present		
PCP Concentration Unit	String	Condition of the measurement, unit for concentration	Not present		
PCP Reference	Text	Literature source for respective PCP	Present or not (null) as data 'FLAMMABILITY(FLASH POINT)' under 'PHYSICAL CHEMICAL DATA'	Reference number in bracket e.g. [017,058,205,275], which appear in the final under SOURCES	Extract from 'SOURCES'

**5.11. pH value**

Name	Type	Description	Present In Source	Format	Conversion Details
CAS (unique identifier)	Integer	The Chemical Abstract Service (CAS) number. This is an integer without hyphens or spaces	Present or not (null) as data 'CAS NUMBER' under 'IDENTIFIERS'	xxxxxx-xx-x e.g. 100-00-5  Digits with two hyphens in the digit. The last is always one digit, the middle two digits, and the first is between two and six digits.	xxxxxxxxx e.g. 100005  Without the two hyphens. The number consists of five to nine digits.
PCP Name	String	pH	Present or not (null) as data 'OTHER PHYSICLA DATA' under 'PHYSICAL CHEMICAL DATA'	String e.g. pH	No change
PCP Value	Float	Value measured for pH value	Present or not (null) as data 'OTHER PHYSICLA DATA' under 'PHYSICAL CHEMICAL DATA'	Float e.g. 8.0	No change
PCP Unit	String	Unit of measured PCP	Not present	No unit	

PCP Method	String	Method of measurement	Not present		Not present
PCP Temp	Float	Condition of the measurement	Not present		
PCP Temp Unit	String	Condition of the measurement, unit for temperature	Not present		
PCP Pressure	Float	Condition of the measurement, pressure	Not present		
PCP Pressure Unit	String	Condition of the measurement, unit for pressure	Not present		
PCP solvent	String	Condition of the measurement solvent	Not present		
PCP Concentration	Float	Condition of the measurement, concentration	Not present		
PCP Concentration Unit	String	Condition of the measurement, unit for concentration	Not present		
PCP Reference	Text	Literature source for respective PCP	Present or not (null) as data 'OTHER PHYSICLA DATA' under 'PHYSICAL CHEMICAL DATA'	Reference number in bracket e.g. [017,058,205,275], which appear in the final under SOURCES	Extract from 'SOURCES'

## 5.12. Viscosity

Name	Type	Description	Present In Source	Format	Conversion Details
CAS (unique identifier)	Integer	The Chemical Abstract Service (CAS) number. This is an integer without hyphens or spaces	Present or not (null) as data 'CAS NUMBER' under 'IDENTIFIERS'	xxxxxx-xx-x e.g. 100-00-5  Digits with two hyphens in the digit. The last is always one digit, the middle two digits, and the first is between two and six digits.	xxxxxxxxx e.g. 100005  Without the two hyphens. The number consists of five to nine digits.
PCP Name	String	Viscosity	Present or not (null) as data 'OTHER PHYSICLA DATA' under 'PHYSICAL CHEMICAL DATA'	String e.g. Viscosity	No change
PCP Value	Float	Value measured for Viscosity	Present or not (null) as data 'OTHER PHYSICLA DATA' under 'PHYSICAL CHEMICAL DATA'	Float number e.g. 0.64	No change

PCP Unit	String	Unit of measured PCP	Present or not (null) as data 'OTHER PHYSICLA DATA' under 'PHYSICAL CHEMICAL DATA'	string e.g. centipoise	No change
PCP Method	String	Method of measurement	Not present		Not present
PCP Temp	Float	Condition of the measurement	Present or not (null) as data 'OTHER PHYSICLA DATA' under 'PHYSICAL CHEMICAL DATA'	Float e.g. 25	No change
PCP Temp Unit	String	Condition of the measurement, unit for temperature	Present or not (null) as data 'OTHER PHYSICLA DATA' under 'PHYSICAL CHEMICAL DATA'	C or c	C
PCP Pressure	Float	Condition of the measurement, pressure	Not present		
PCP Pressure Unit	String	Condition of the measurement, unit for pressure	Not present		
PCP solvent	String	Condition of the measurement solvent	Not present		
PCP Concentration	Float	Condition of the measurement, concentration	Not present		
PCP Concentration Unit	String	Condition of the measurement, unit for concentration	Not present		

PCP Reference	Text	Literature source for respective PCP	Present or not (null) as data 'OTHER PHYSICLA DATA' under 'PHYSICAL CHEMICAL DATA'	Reference number in bracket e.g. [017,058,205,275], which appear in the final under SOURCES	Extract 'SOURCES' from
---------------	------	--------------------------------------	--	---	------------------------

**5.13. Refractive Index**

Name	Type	Description	Present In Source	Format	Conversion Details
CAS (unique identifier)	Integer	The Chemical Abstract Service (CAS) number. This is an integer without hyphens or spaces	Present or not (null) as data 'CAS NUMBER' under 'IDENTIFIERS'	xxxxxx-xx-x e.g. 100-00-5  Digits with two hyphens in the digit. The last is always one digit, the middle two digits, and the first is between two and six digits.	xxxxxxxxx e.g. 100005  Without the two hyphens. The number consists of five to nine digits.

PCP Name	String	Refractive index	Present or not (null) as data 'OTHER PHYSICLA DATA' under 'PHYSICAL CHEMICAL DATA'	String e.g. Refractive index	No change
PCP Value	Float	Value measured for Refractive index	Present or not (null) as data 'OTHER PHYSICLA DATA' under 'PHYSICAL CHEMICAL DATA'	Float number e.g. 1.49594	No change
PCP Unit	String	Unit of measured PCP	No present	No unit	
PCP Method	String	Method of measurement	Not present		Not present
PCP Temp	Float	Condition of the measurement	Present or not (null) as data 'OTHER PHYSICLA DATA' under 'PHYSICAL CHEMICAL DATA'	Float e.g. 25	No change
PCP Temp Unit	String	Condition of the measurement, unit for temperature	Present or not (null) as data 'OTHER PHYSICLA DATA' under 'PHYSICAL CHEMICAL DATA'	C or c	C
PCP Pressure	Float	Condition of the measurement, pressure	Not present		
PCP Pressure Unit	String	Condition of the measurement, unit for pressure	Not present		

PCP solvent	String	Condition of the measurement solvent	Not present		
PCP Concentration	Float	Condition of the measurement, concentration	Not present		
PCP Concentration Unit	String	Condition of the measurement, unit for concentration	Not present		
PCP Reference	Text	Literature source for respective PCP	Present or not (null) as data 'OTHER PHYSICLA DATA' under 'PHYSICAL CHEMICAL DATA'	Reference number in bracket e.g. [017,058,205,275], which appear in the final under SOURCES	Extract from 'SOURCES'

**5.14. Octanol Water Partition Coefficient**



Name	Type	Description	Present In Source	Format	Conversion Details
CAS (unique identifier)	Integer	The Chemical Abstract Service (CAS) number. This is an integer without hyphens or spaces	Present or not (null) as data 'CAS NUMBER' under 'IDENTIFIERS'	xxxxxx-xx-x e.g. 100-00-5  Digits with two hyphens in the digit. The last is always one digit, the middle two digits, and the first is between two and six digits.	xxxxxxxxx e.g. 100005  Without the two hyphens. The number consists of five to nine digits.
PCP Name	String	Log octanol/water partition coefficient	Present or not (null) as data 'OTHER PHYSICLA DATA' under 'PHYSICAL CHEMICAL DATA'	String e.g. Octanol/water partition coefficient or Log octanol/water partition coefficient	Change to Log octanol/water partition coefficient
PCP Value	Float	Value measured for Octanol/water partition coefficient	Present or not (null) as data 'OTHER PHYSICLA DATA' under 'PHYSICAL CHEMICAL DATA'	Float number e.g. 1.3	No change

PCP Unit	String	Unit of measured PCP	Not present	No unit	
PCP Method	String	Method of measurement	Not present		Not present
PCP Temp	Float	Condition of the measurement	Not present		
PCP Temp Unit	String	Condition of the measurement, unit for temperature	Not present		
PCP Pressure	Float	Condition of the measurement, pressure	Not present		
PCP Pressure Unit	String	Condition of the measurement, unit for pressure	Not present		
PCP solvent	String	Condition of the measurement solvent	Not present		
PCP Concentration	Float	Condition of the measurement, concentration	Not present		
PCP Concentration Unit	String	Condition of the measurement, unit for concentration	Not present		
PCP Reference	Text	Literature source for respective PCP	Present or not (null) as data 'OTHER PHYSICLA DATA' under 'PHYSICAL CHEMICAL DATA'	Reference number in bracket e.g. [017,058,205,275], which appear in the final under SOURCES	Extract from 'SOURCES'

### **5.15. Partition Coefficient**

Not sure what Partition Coefficient is.

## 6. ADME Related Properties and Descriptors

Name	Type	Description	Present In Source	Format	Conversion Details
CAS (unique identifier)	Integer	The Chemical Abstract Service (CAS) number. This is an integer without hyphens or spaces	Present or not (null) as data 'CAS NUMBER' under 'IDENTIFIERS'	xxxxxx-xx-x e.g. 100-00-5 Digits with two hyphens in the digit. The last is always one digit, the middle two digits, and the first is between two and six digits.	xxxxxxxxx e.g. 100005 Without the two hyphens. The number consists of five to nine digits.
1 <sup>st</sup> acidic pK <sub>a</sub>	Float	Strongest acidic dissociation constant	Not present		
2 <sup>nd</sup> acidic pK <sub>a</sub>	Float	2 <sup>nd</sup> strongest acidic dissociation constant	Not present		
3 <sup>rd</sup> acidic pK <sub>a</sub>	Float	3 <sup>rd</sup> strongest acidic dissociation constant	Not present		
4 <sup>th</sup> acidic pK <sub>a</sub>	Float	4 <sup>th</sup> strongest acidic dissociation constant	Not present		
1 <sup>st</sup> basic pK <sub>a</sub>	Float	Strongest basic dissociation constant	Not present		
2 <sup>nd</sup> basic pK <sub>a</sub>	Float	2 <sup>nd</sup> strongest basic dissociation constant	Not present		
3 <sup>rd</sup> basic pK <sub>a</sub>	Float	3 <sup>rd</sup> strongest basic dissociation constant	Not present		

4 <sup>th</sup> basic pK <sub>a</sub>	Float	4 <sup>th</sup> strongest basic dissociation constant	Not present		
LogP	Float	Octanol/water partition coefficient of neutral form	See 5.14	See 5.14	see 5.14
LogD at pH=0	Float	Octanol/water distribution coefficient at pH=0	Not present		
LogD at pH=1	Float	Octanol/water distribution coefficient at pH=1	Not present		
LogD at pH=2	Float	Octanol/water distribution coefficient at pH=2	Not present		
LogD at pH=3	Float	Octanol/water distribution coefficient at pH=3	Not present		
LogD at pH=4	Float	Octanol/water distribution coefficient at pH=4	Not present		
LogD at pH=5	Float	Octanol/water distribution coefficient at pH=5	Not present		
LogD at pH=6	Float	Octanol/water distribution coefficient at pH=6	Not present		
LogD at pH=7	Float	Octanol/water distribution coefficient at pH=7	Not present		
LogD at pH=7.4	Float	Octanol/water distribution coefficient at pH=7.4	Not present		
LogD at pH=8	Float	Octanol/water distribution coefficient at pH=8	Not present		
LogD at pH=9	Float	Octanol/water distribution coefficient at pH=9	Not present		
LogD at pH=10	Float	Octanol/water distribution coefficient at pH=10	Not present		
LogD at pH=11	Float	Octanol/water distribution coefficient at pH=11	Not present		
LogD at pH=12	Float	Octanol/water distribution coefficient at pH=12	Not present		
LogD at pH=13	Float	Octanol/water distribution coefficient at pH=13	Not present		
LogD at pH=14	Float	Octanol/water distribution coefficient at pH=14	Not present		
MR	Float	Molar refractivity	Not present		
HBD	Float	Number of Hydrogen-bond donors	Not present		
HBA	Float	Number of Hydrogen-bond acceptors	Not present		
PSA	Float	Polar Surface Area	Not present		

Doc. Identifier:

OpenMolGRID-1-D1.1d-01042-1-  
NTPDataSpec

**NTP DATA SPEC**

Date: 07/01/2004

---

TPSA	Float	Topological Polar Surface Area	Not present		
------	-------	--------------------------------	-------------	--	--

## 7. References

- [1] D. McCourt, "Data Warehouse software Specification," <https://hermes.chem.ut.ee/bscw/bscw.cgi/d2137/OpenMolGRID-12-D1.1a-0101-3-0-MOLDW>, 10/09/03.
- [2] E. Benfenati and A. Papp, "Properties and priorities of the data for pharmaceutical and phytopharmaceutical compounds," <https://hermes.chem.ut.ee/bscw/bscw.cgi/d6500/OpenMolGRID-12-D1.3a-0108-1-0>, 10/09/03.
- [3] D. McCourt, J. Jing and W. Dubitzky, "Description of NTP," <https://hermes.chem.ut.ee/bscw/bscw.cgi/d6202/OpenMolGRID-1-D1.4b-0110-1-0>, 12/09/03.