Work Package 2 "Identification of relevant target substances in BREFs" of the HAZBREF project funded by Interreg Baltic Sea Region

Report

Approaches for a better use of available data to prevent or reduce releases of substances of concern from industrial installations

Annex 4

HAZBREF
Hazardous industrial chemicals in the IED BREFs

Interreg
Baltic Sea Region

EUROPEAN UNION
EUROPEAN REGIONAL DEVELOPMENT FUND

Identification of chemicals belonging to chemical classes used as textile auxiliaries

A study on behalf of the German Environment Agency
in the context of HAZBREF

Dr. Barbara Werschkun
Wissenschaftsbüro • Berlin

2 April 2020

CONTENTS

1. Starting point

This study is performed under a contract with the German Environment Agency (UBA) as part of its engagement in the EU Interreg project HAZBREF (Hazardous Industrial Chemicals in the IED BREFs). The overall objective of UBA in HAZBREF is to identify relevant target substances for selected industry sectors, in particular with regard to industrial wastewater. The specific objective of this contract study is to compile lists of chemical substances according to their chemical structure, starting from a list of chemical classes that are frequently used as auxiliaries in the textile industry. This list had been compiled by industry representatives as their contribution to the HAZBREF project [file *TXT-BREF-Quest_Chemicals_cond_(08-2018).xlsx*]. The aim was not to obtain comprehensive lists of all substances belonging to a certain substance class, but rather a set of representative substances that would be sufficient for an evaluation of the physico-chemical and environmental properties of a given substance class. These evaluations were not part of the present study.

At the start of this study, the above-mentioned list of chemical classes was modified and further expanded to match the structuring of chemical groups according to technical functions as provided in Annex 8 of the current BREF for the textile industry (European Commission 2003). The resulting document (file *Overview chemical classes.xlsx*) served as the "master list of chemical classes" for the further development of the study. The master list contains more than 200 entries of chemical classes under the following process headings:

1 Surfactants
2 Auxiliaries and finishing agents for fibre and yarn manufacturing
3 Sizing agents
4 Detergents/wetting agents
5 Auxiliaries containing sequestering agents
6 Dyeing auxiliaries
7 Printing auxiliaries
8 Finishing auxiliaries
9 Coating compounds and auxiliaries
10 Solvents
11 Other chemicals

The process headings are further subdivided into technical functions, e.g.

7 Printing auxiliaries
- *7A Thickening agents*
- *7B Binders*
- *7C Fixing agents*
- *7D Plasticizers*
- *7E Emulsifiers*

Some technical functions are needed for different processes, e.g. emulsifiers are used during fibre and yarn manufacturing as well as during printing. Some classes of chemicals are also used for more than one technical function, e.g. surfactants are used as emulsifiers, detergents, wetting agents, dispersing and levelling agents during printing etc. Therefore, several entries in the master lists are duplicates. Several others refer to individual substances, e.g. magnesium chloride or perchloroethylene, in which case no further search is needed. Some substance classes are described in terms that do not allow further searches for representative substances – either because the

description is too general, e.g. “fluorine derivatives”, or because the wording is not suitable for database entry, e.g. “resin-based repellents (condensation products of fatty acids, alcohols or amines with methylolated melamines)”. In these cases, more information from within the textile industry would be helpful – regarding the chemical structure or even regarding typical substances used.

Following the establishment of a structure-based approach to identify chemicals corresponding to defined substance classes used in the textile industry, the second objective of this study was the application of this approach to other industry sectors. For these industries, no systematic compilations of chemical substance classes were available. Instead, lists of individual chemicals were provided from case studies. The task was, first, to define appropriate substance classes based on these chemicals and then to apply the structure-based approach for the identification of further chemicals that belong to those classes. Case studies were available for the following industry sectors: production of polymers, surface treatment of metals, and manufacture of fertilizers.

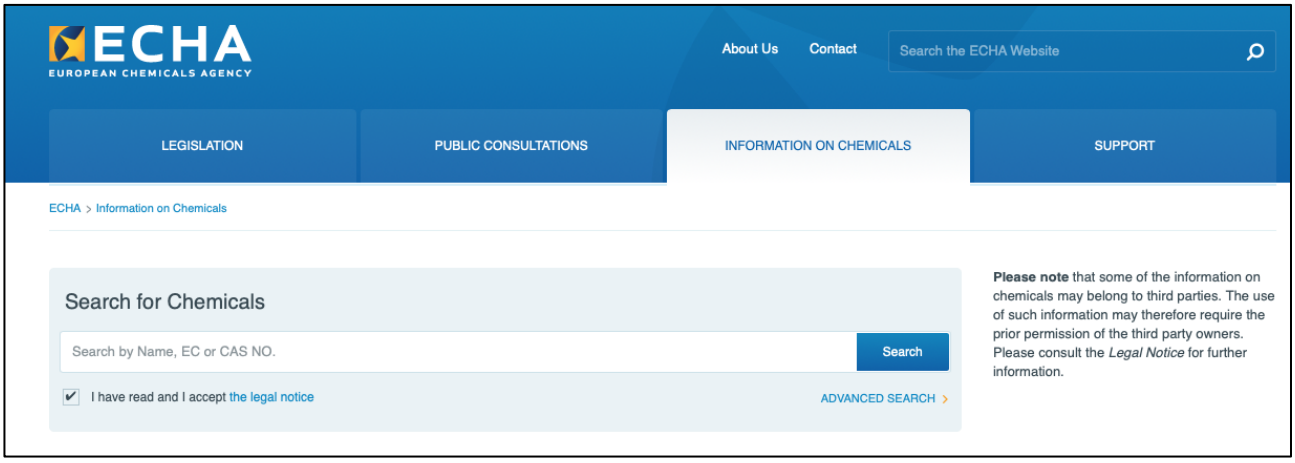
2. Procedure

All searches are performed in the database of chemicals maintained by the European Chemicals Agency (ECHA) which is publicly available online (https://echa.europa.eu/information-on-chemicals). This database is generated from information submitted by industry under Regulation (EC) No 1907/2006 concerning the Registration, Evaluation, Authorisation and Restriction of Chemicals (REACH) and Regulation (EC) No 1272/2008 on classification, labelling and packaging of substances and mixtures (CLP) and comprises a total of almost 250,000 entries. However, only about 22,000 of these entries refer to substances that are registered under REACH, along with the legally required information regarding substance identity, inherent properties, uses and exposure. As these information packages were compiled by a multitude of different registrants, the completeness and quality of the data available for individual substances is very heterogenous. This seriously limits the applicability of the database for the compilation of comprehensive lists of substances with certain characteristics. For instance, in the beginning of the HAZBREF project, the UBA searched the database for all substances registered for the Sector of Use 5, "Manufacture of textiles, leather, fur", which resulted in >900 matches. However, consultation with industry representatives revealed that, on the one hand, these results contained substances that are not actually used in the textile industry while, on the other hand, substances that are used were missing from this list. Other examples for the inconsistency of the ECHA database will be discussed later.

2.1 Search by common name

As a first step, the common name of the substance class as included in the "master list" (see chapter 1) is typed into the search field "Search by Name, EC or CAS No." in the "Simple search" function of the ECHA online database (Figure 1). If there is more than one common name in use for the same substance class, e.g. 'fatty alcohol ethoxylates', 'ethoxylated fatty alcohols' and 'alcohol polyglycol ethers', the search is repeated accordingly.

Figure 1. The 'Simple search' window in the ECHA database of chemical substances.

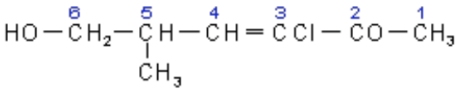


In most cases, this already generates a number of results, which are good representatives of the substance class in question. If the obtained list is not considered sufficient, the individual entries can serve as starting points for the derivation of a generic systematic name fragment to be used in additional searches as discussed below.

2.2 Search by systematic name fragment

Internationally agreed rules for the systematic nomenclature of organic and inorganic chemicals are published by the International Union for Pure and Applied Chemistry (IUPAC 2019). With some adaptations, this system is also applied for indexing by the Chemical Abstracts Service (CAS 2007). The preferred type of nomenclature used in the IUPAC system is the so-called "substitutive nomenclature". Simply put, this involves the identification and naming of the principal molecular skeleton, the determination of appropriate prefixes to reflect additional functional groups or modifications, and the assembly of all name components into a complete name. An example is given in Figure 2.

Figure 2. Example for applying substitutive nomenclature according to IUPAC rules.



Principal group:	–CO–	-one
Principal chain:	C–C–C–C–C–C	Hexane
Principal chain including principal group:	C–C–C–C–CO–C	2-Hexanone
Subtractive modification:	C–C–C=C–CO–C	3-Hexen-2-one
Prefixes:	Cl–	Chloro-
	HO–	Hydroxy-
	CH_3–	Methyl-

Together with later rules this leads to the name:
3-Chloro-6-hydroxy-5-methyl-3-hexen-2-one

The molecule in this example contains two functional groups, a keto group and a hydroxy group. In order to choose the principal group for naming, the IUPAC system has established hierarchy rules. According to these, ketones generally take precedence over alcohols. However, if the naming rules are not followed correctly, the same molecule might also be identified as 4-chloro-2-methyl-5-oxo-3-hexenol. Depending on the complexity of the structure and on the expertise of the person responsible for the naming, it cannot be excluded that alternative names, even though formally 'incorrect', may be entered into registers and databases, in particular when these are compiled from many different entry points, such as the ECHA database.

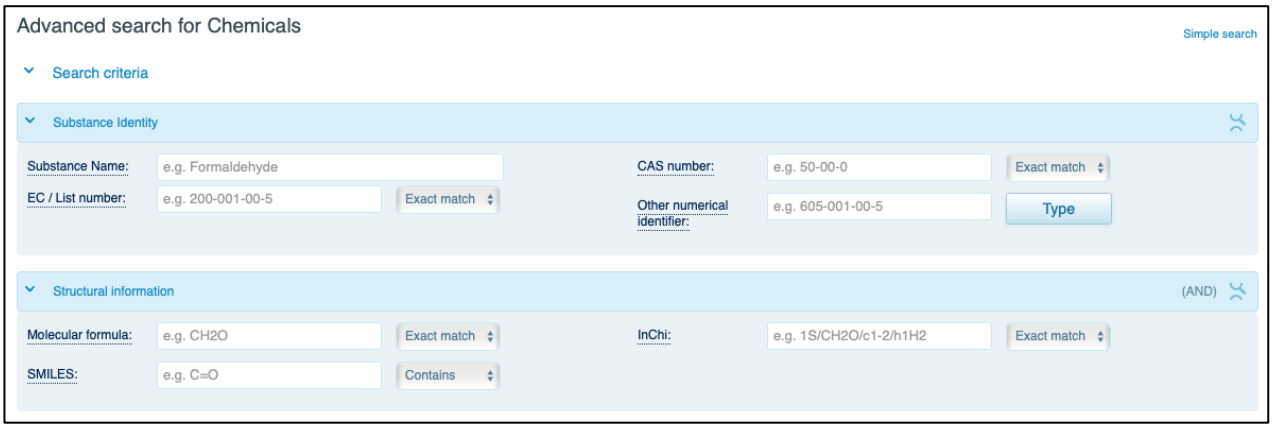
In general, carbon chains and rings are named according to their length based on Greek numerals (in the example above 'hexa' is used to name a 6-carbon chain). However, as an exception to this systematic approach, the IUPAC system also retains a considerable number of trivial and semi-systematic names that have traditionally been used for naming organic compounds. Examples include many cyclic structures such as benzene, naphthalene, pyridine etc., prefixes for short-chain acyclic substituents, such as vinyl, isopropyl, acetyl etc., several carboxylic acids derived from natural products, such as malonic acid, glutaric acid, oleic acid etc. and many more. All these additional features need to be taken into account when searching for a possible name of a given substance. For instance, the substance with the molecular formula CH_3-$[CH_2]_{16}$-COO-$CH(CH_3)_2$ might be represented by any of the following names: octadecanoic acid 2-methylethyl ester; 2-methylethyl octadecanoate; octadecanoic acid isopropyl ester; isopropyl octadecanoate; stearic acid 2-methylethyl ester; 2-methylethyl stearate; stearic acid isopropyl ester; isopropyl stearate. Entries in the ECHA database frequently contain a list of alternative names. Thus, many substances may be found by more than one specific name.

Rather than determining the correct name – and possible alternatives – for one specific substance, the challenge of this study lies in formulating a name fragment that may serve for the identification of many similar substances. This is not the same as formulating a generic name that would cover a group of substances. For example, the generic name “alkanol”, used for a simple search in the ECHA database, results in only 29 hits – all relating to derivatives or reaction products of alkanols, but not to any specific substances that actually *are* alkanols. The names of some specific substances may allow to deduce a common name fragment, e.g. the names for some alcohols such as methanol, ethanol, propanol, butanol, pentanol etc. all share the common fragment “anol”. If this fragment is used for a simple search in the ECHA database, it results in more than 12,000 hits. Thus, finding a name fragment with an adequate degree of specificity is a critical step in this study. In some cases, this problem may be alleviated by combining a rather general name fragment, which by itself produces too many hits, with additional search parameters, such as SMILES code fragments, sector of use etc (see below).

2.3 Search by SMILES code fragment

To allow better processing of chemical structure descriptors by software, line notation systems were created. The Simplified Molecular-Input Line-Entry System (SMILES) was developed by the company Daylight Chemical Information Systems (Daylight 2011). Atoms in a molecule are represented by the standard abbreviation of the chemical element (hydrogen atoms in organic molecules are usually omitted). Specific bond types are represented by symbols, e.g. “-“ for a single bond (which is usually omitted) or “=” for a double bond. For example, 4-hexene-1-ol, CH_3-CH=CH-CH_2-CH_2-CH_2-OH, is represented by the SMILES notation CC=CCCCO. Sidechains branching off the main chain are put in brackets, for example CC(=O)O for acetic acid. Since there can be ambiguity over what is considered the main chain and what is considered a sidechain, it is possible to represent a single molecule by more than one SMILES code (this will be illustrated with several examples in chapter 3). Algorithms have been developed to generate unequivocal, so-called “canonical” SMILES strings for each chemical structure. This study uses a SMILES generator referenced in the ECHA guidance that is freely available online: http://www.cheminfo.org/flavor/malaria/Utilities/SMILES_generator___checker/index.html.

Figure 3. The sub-menu “Structural information” in the “Advanced search” function of the ECHA database.



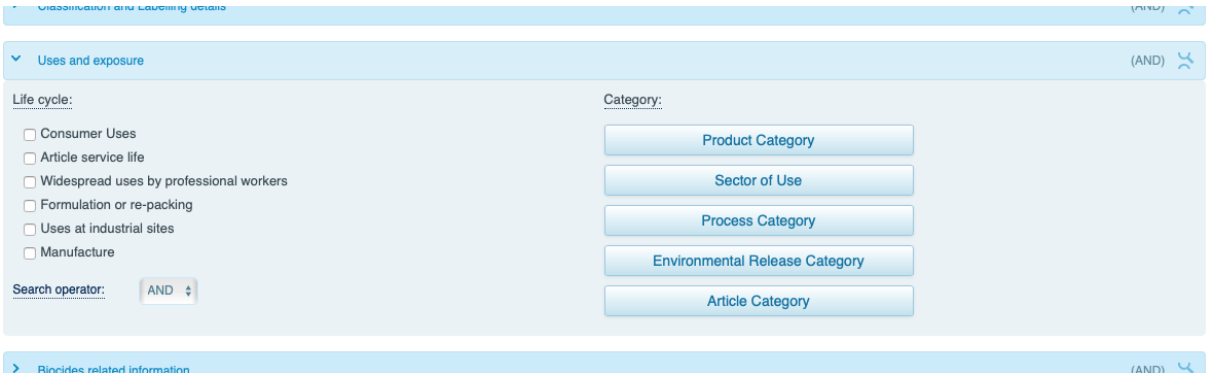
Instead of defining a complete SMILES code for one molecule, the search for substances within chemical substance classes requires the definition of a SMILES code fragment, which represents

the typical structural element of this substance class. For instance, the fragment COC is present in all molecules that contain an ether bond. The fragment C(=O)OC describes an organic ester. A search by SMILES code is performed within the "Advanced search" function of the ECHA database by using the function "SMILES contains …" in the sub-menu "Structural information" (Figure 3).

2.4 Filtering and processing of search results

In some cases, the search strategies described above lead to a large number of substances which cannot be further narrowed down on structure-related information alone. In those cases, other parameters from the ECHA "Advanced search" menu were added (Figure 4), for instance "Regulatory context" (in particular the Biocidal Product Regulation) or "Uses and exposures" (in particular sector of use 5 – manufacture of textiles, leather, fur). Other parameters which are available in the "Advanced search" function, even though they were not used in this study, relate to "Properties of concern" or "Classification and labelling details".

Figure 4. The sub-menu "Uses and exposure" in the "Advanced search" function of the ECHA database.



Search results obtained from the ECHA database are downloaded as Excel files for further processing. The downloaded files contain the substance name, EC number, and CAS number, if available. Processing includes the following steps:

- Verification if the obtained substances really represent the desired substance class
- Exclusion of substances that do not seem representative for the substance class in this context, for instance because they contain additional functional groups that would interfere with the technical function under consideration (depending on the situation, this is either performed manually or by applying the Excel auto filter)
- Deletion of duplicates
- Grouping into sub-classes

Not all of these steps are applicable in all cases. Examples and further explanations are provided in chapter 3. In general, the processing of search results relies heavily on expert judgment and hence requires a solid background knowledge of organic chemistry.

As already pointed out in chapter 1, completeness was not the aim of the performed searches. Therefore, in many cases the searches were stopped once a sufficient number of representative results could be obtained.

3. Results and discussion

3.1 Surfactants

1A. Non-ionic

1A.1 Fatty alcohol ethoxylates

These are polyethers consisting of an aliphatic alcohol (chain length C8-C24) and up to 40 ethylene glycol units, with a terminal free hydroxyl group. The generic structure can be described as $CH_3(CH_2)_{6\text{-}22}CH_2(OCH_2CH_2)_{1\text{-}40}OH$ (which is not entirely correct as the fatty alcohol chain may also contain double bonds or sidechains). An initial search using the generic term “fatty alcohol ethoxylate” already gives 12 results. Four of these are excluded because they contain additional functional groups that are assumed to significantly modify the chemical characteristics of the compounds (in this case: sulphate, dimethylpolysiloxane, and a terminal butyl ether instead of a free hydroxyl group). Another entry is excluded because it appears to refer to a reaction mixture rather than a single compound. A search for the characteristic SMILES code fragment CCCCCCCCOCCO gave 80 results, but the majority (58) of them were excluded because of additional functional groups. Name fragments for additional searches can be derived from the systematic names of the initially retrieved results. If too many results are obtained, the search needs to be further refined. The fragment “oxy)ethoxy” gave more than 600 results. The search was further refined by adding the category “Sector of use = SU 5: Manufacture of textile, leather, fur” in the sub-menu “Uses and exposure” in the ECHA database “Advanced search” function, thereby reducing the number of results to 9 (of which 8 were manually excluded). The fragment “oxy-1,2-ethanediyl” gave almost 900 results. These were further refined by using the Excel “auto filter” function and removing all entries that contain additional functional groups, which are assumed not to be relevant for substances used as surfactants (in this case: amin, imin, amid, cyan, nitril, phosph, sulph/sulf, thio, phen, benz, acid, oate, carbox, oxo, chlor, sil, polymer, reaction). The fragment “-ol, ethoxylated” gave 107 results, which were filtered manually.

Combining all searches and removing duplicates resulted in 143 individual substance entries, which can all be considered representatives of the chemical substance class “fatty alcohol ethoxylates”. However, only 48 of these substances have CAS numbers. The substances not having CAS numbers appear to be specific derivatives of a more generic substance identity, as shown in Table 1 for the example of ethoxylated decanol.

Table 1. Derivatives of ethoxylated decanol found in the ECHA database.

Substance Name	EC Number	CAS Number
Decan-1-ol, ethoxylated	500-046-6	26183-52-8
Decan-1-ol, ethoxylated (>=3-EO) (26183-52-8)	931-713-3	-
Decan-1-ol, ethoxylated (>2.5 moles EO) (CAS: 26183-52-8)	931-960-7	-
Decan-1-ol, ethoxylated (3EO) (26183-52-8)	935-102-2	-
Decan-1-ol, ethoxylated (4-EO) (26183-52-8)	931-757-3	-
Decan-1-ol, ethoxylated (5-EO) (26183-52-8)	931-758-9	-
Decan-1-ol, ethoxylated (6EO) (26183-52-8)	935-101-7	-
Decan-1-ol, ethoxylated (EO = 5)	939-521-1	-
Decan-1-ol, ethoxylated (polymer)	938-595-2	-

1A.2 Fatty acid ethoxylates

These are structurally quite similar to the fatty alcohol ethoxylates discussed above. Instead of the fatty alcohol unit, they contain a fatty acid unit linked by an ester bond to a chain of ethylene glycol units, with a terminal free hydroxyl group. The generic structure can be described as $CH_3(CH_2)_{6-22}CO(OCH_2CH_2)_{1-40}OH$ (again, not strictly correct as the fatty acid may also contain double bonds or sidechains).

The generic name search and the search by SMILES code fragment are performed in analogy to the fatty alcohol ethoxylates. The search for systematic name fragments was restricted to the term "acid, ethoxylated". No further searches were performed, because the combination of these three steps already gave a sufficient number of results (28 individual substances with CAS numbers).

1A.3 Alkylphenol ethoxylates (APEOs)

These substances consist of a chain of 1-40 ethylene glycol units linked to an alkyl phenol: $R\text{-}C_6H_4\text{-}O(CH_2CH_2O)_{1-40}H$ (R = alkyl). Queries were performed as described above, using the generic name, a characteristic SMILES code fragment (CC1=CC=C(OCCO)C=C1) and two systematic name fragments ("phenol, ethoxylated" and "ylphenoxy)eth"). Noticeably, the search by SMILES code fragment could have been omitted, since it only gave very few results, which were all duplicates. This is quite different from the previous substance classes. This may be the case, because it is more difficult to denote cyclic substructures (i.e. the phenyl ring) with SMILES than linear ones. Again, the search was terminated after a sufficient list of representative substances had been obtained.

1A.4 Fatty amine ethoxylates

These are structurally similar to fatty alcohol ethoxylates and fatty acid ethoxylates, with a polyethylene glycol sidechain linked to a fatty amine, $CH_3(CH_2)_{6-22}CH_2NR(CH_2CH_2O)_{1-40}H$, where R can be alkyl or hydrogen. The generic name search and the search by SMILES code fragment are performed in analogy to the fatty alcohol and fatty acid ethoxylates. The chosen systematic name fragment was typed in two different ways: "amine ethoxylated" (without a comma) and "amine, ethoxylated" (with a comma). In both cases, quite different search results were obtained, pointing out an apparent sensitivity of the ECHA database to interpunctuation.

1A.5 Triglyceride ethoxylates

These substances are derived from glycerol (1,2,3-propanetriol), in which all three hydroxyl functions are linked to polyethylene glycol chains, the terminal hydroxy group of which is esterified with a fatty acid unit. Due to the complexity and variability of the structure, it cannot be easily described by a generic molecular formula. For the same reason, a characteristic SMILES code fragment cannot be defined. The simple search using the generic common name "triglyceride ethoxylate" does not lead to any result. A search using the name fragment "triglyceride" leads to 37 entries, but manual processing leaves only one valid entry (because the other ones do not contain any ethoxylate groups). A search for simply "glyceride" results in several hundred entries, which are then further processed by the Excel auto filter (contains "ethoxy" AND does not contain "reaction mass") and manually validated (under removal of all pure mono- or diglycerides). This way, eleven representative substances can be retrieved.

1A.6 Ethylene oxide/propylene oxide adducts

For this substance class, further clarification is needed. As the term "ethylene oxide adduct" is synonymous to the term "ethoxylate", there seems to be an overlap with some of the previously

discussed substance classes. From the generic name, it is not clear, to which other component the "ethylene oxide" or "propylene oxide" unit is attached.

1B. Anionic

1B.1 Alkyl sulphonates

These substances are derived from sulphonic acid, substituted with an alkyl chain: R-SO_2-O^-. Generic name searches with various spellings generate very few results. Results obtained by SMILES code search are all judged invalid by manual inspection. Eventually, the search for "sulfonic acid" combined with the Excel auto filter (contains "alk" AND does not contain "chloro") results in a suitable list of representative substances, which are further divided into sub-groups according to the structure of the alkyl chain (saturated, unsaturated, hydroxylated) and sorted by chain length. (Both the chain length and the presence or absence of hydroxy groups can be assumed to influence the environmental characteristics of the substances.)

1B.2 Alkylaryl sulphonates

In this substance group, a sulphonic acid unit is attached to a benzene ring, which is substituted with an alkyl chain: R-C_6H_4-SO_2-O^-. The search by generic name in different spellings gives only one valid result, but a search for the name fragment "ylbenzenesulfonate" followed by Excel auto filter (does not contain azo, naphth, amin, reaction, pyrid, piperid, quinol, nitro, phenyl) results in a suitable list of representative substances, which are divided into subgroups according to the the structure of the alkylaryl unit (dodecylbenzensulfonates, other long-chain alkyl substituents, short-chain alkyl substituents), which can be assumed to influence properties such as water solubility or degradability.

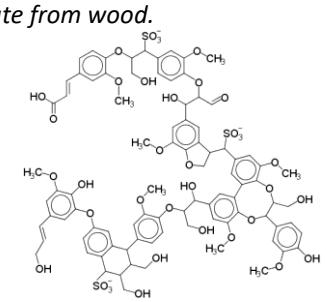
1B.3 Naphthalenesulphonates

These substances are derived from a naphthalene ring that contains a sulfonate group and an alkyl substituent: R-$C_{10}H_6$-SO_2-O^-. Queries using the generic name in different spellings result in a large number of hits, which are further refined by Excel auto filter and then grouped according to the nature of their substituents (unsubstituted; hydroxyl; short-chain alkyl; longer chain alkyl; dimeric structures; miscellaneous). More information regarding specific substances used as surfactants in the textile industry would be helpful to make a decision on whether or not all these different subgroups are relevant in this context.

1B.4 Ligninsulphonates

These substances are derived from lignin, a class of complex organic polymers naturally present in plants and consisting of cross-linked hydroxyphenylalkanes (Figure 5). Due to the size and complexity of these molecules, it is not possible to define a characteristic SMILES code fragment. Queries using the generic name (and name fragments) in different spellings result in a suitable list of representative substances.

Figure 5. Structural model of ligninsulphonate from wood.



1B.5 Alkyl sulphates

These substances are derived from sulfuric acid, esterified with fatty alcohols: CH_3-$(CH_2)_n$-CH_2-O-SO_2O^-. Initial searches using the generic name give only a small number of results. A search using the characteristic SMILES code fragment reveals that the ECHA database contains SMILES codes in a different notation than the one obtained with the SMILES code generator that is recommended in the ECHA guidance (ECHA 2017). While the latter is COS(=O)(=O)[O-], only the use of the alternative version COS([O-])(=O)=O gives a list of results, which are further refined both manually and by Excel auto filter and sorted into sub-groups according to the structure of the alkyl chain (linear or branched) and the identity of the cation (e.g. sodium, potassium or ammonium).

1B.6 Alcohol ethoxysulphates

This substance class is also derived from sulfuric acid and contains fatty alkyl groups, but instead of being directly linked, as in the alkyl sulphates discussed above, they are separated by an unspecified number of ethoxy groups: CH_3-$(CH_2)_m$-CH_2-$(OCH_2CH_2)_n$-O-SO_2O^-. Initial searches using the generic name in alternative spellings give no results. The SMILES code fragment used above to identify alkyl sulphates is amended by addition of an ethoxy group to CO**CCO**S([O-])(=O)=O, giving a number of representative substances. From the systematic names of these initial results, the name fragment "poly(oxy-1,2-ethanediyl), alpha.-sulfo-.omega.-hydroxy-," is identified for an additional search. The results are grouped according to chain length, number of ethoxy groups, and cations, as far as it was feasible.

1B.7 Sulphated alkanolamides

These substances are derivatives of fatty acid amides - $CH_3(CH_2)_nCONRR'OSO_2O^-$ - in which the nitrogen atom of the amide group is attached to at least one alkyl chain carrying a sulphate group. A simple example is sulphated N-(hydroxyethyl)-octylamide, $CH_3(CH_2)_6$**CON**HCH_2CH_2**OSO_2O^-**. As the two defining structural elements, the amide group (CON) and the sulphate group (SO_2O^-), may be linked through an alkyl chain of any length, it is not possible to define a typical SMILES code fragment for this substance class. Searches were therefore limited to various combinations of the name fragments sulfate/sulphate and amid*, followed by the application of Excel auto filter as well as manual refinement.

1B.8 Dialkylsulphosuccinates

These are dialkyl esters of succinic acid, $ROOCCH_2CH_2COOR'$, in which one of the methylene groups is attached to a sulphonyl group, -SO_2O^-, to give the generic structure $ROOCCH(SO_2O^-)CH_2COOR'$, where R and R' may either be the same or different. The initial search using part of the trivial name in alternative spellings (sulfosuccinate/sulphosuccinate) already gives around twenty representative substances.

1B.9 Sulphated/sulphonated vegetable oils

Vegetable oils are natural components of seeds consisting of fatty acids linked to a triglyceride core. For industrial applications, sulfate or sulfonate groups are added to the fatty acid chain in order to increase its hydrophilicity. It is not possible to define a characteristic SMILES code fragment for this substance class. Searches performed on various combinations of the terms "sulf*/sulph*", "sulfate/sulphate", "sulfonate/sulphonate", and "(vegetable) oil" give a large number of representative results, which can be sorted into groups of sulphates *vs* sulphonates on the

one hand and substances derived from natural vegetable oils, modified vegetable oils or vegetable oil components (i.e. fatty acids), on the other hand.

1B.10 Alkyl carboxylates
These are salts derived from organic acids, $RCOO^-$, and various cations such as sodium, potassium, calcium, ammonium etc. Searches are based on both the generic name and the SMILES code fragment CCCCCC([O-])=O, assuming that a minimum chain length of six carbon atoms is required to achieve surfactant properties. The results are filtered to retrieve mainly salts composed of fatty acid carboxylates and sodium or potassium.

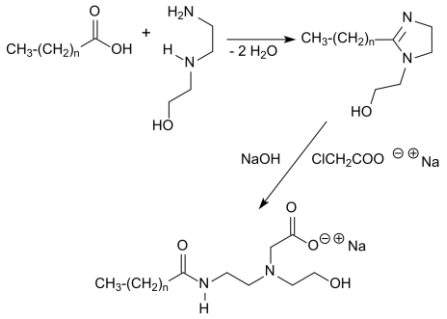
1B.11 Alkylether phosphates
This substance class contains alkyl ethers, ROR', containing an additional hydroxy group, which is linked to a phosphate residue as in the general formula $ROR'OP(OR'')_2O$. In practice, R' most often stands for ethylene, CH_2CH_2, or (poly)ethoxyethylene, $(CH_2CH_2O)_nCH_2CH_2$. This assumption allows the formulation of the characteristic SMILES code fragment CCOCCOP. Additional searches are performed with the name fragments "(alkyl)ether" and "phosphate". As observed earlier, different results were obtained, depending on whether or not the search terms were separated by a comma. The list of representative substances is divided into subgroups according to the structure of the terminal alkyl chain.

1C. Cationic
1C.1 Quaternary ammonium compounds
In quaternary ammonium compounds, four alkyl or aryl substituents are linked to a nitrogen atom, which thereby obtains a permanent positive charge ($R^1R^2R^3R^4N^+$). For the identification of representative substances, this term is too general. Further clarification is needed regarding the nature of the substituents. To some extent, this substance class is covered later under *6C.10 Quaternary ammonium salts with C12-C14 fatty alkyl side chains* and *6C.11 Quaternary ammonium salts with aromatic ring systems*.

1D. Amphoteric
1D.1 Betaine derivatives
Betaines, derived from the natural product betaine (*N,N,N*-trimethylglycine), are zwitter ions that contain a positively charged nitrogen atom and a negatively charged carboxylate group in the same molecule, $R^1R^2R^3N^+-R^4-COO^-$. (Phosphonium betaines, which contain phosphorous instead of nitrogen will be neglected in this context.) In theory, R^4 can be of variable chain length and structure. In practice, it is most often CH_2. This is exemplified by searches performed on a number of possible SMILES code fragments. Results are only retrieved for C[N+](C)(C)**CC**([O-])=O, and not for the alternatives C[N+](C)(C)**C**([O-])=O or C[N+](C)(C)**CCC**([O-])=O. Most hits, however, result from a simple search using the generic term "betaine". All results are sorted into subgroups according to the nature of the alkyl substituents R^1-R^3.

1D.2 Imidazolines
In the actual sense, imidazolines are five-membered cyclic compounds consisting of two nitrogen atoms linked to an ethylene group and separated by a sp^2 hybridized carbon atom. In the context of amphoteric surfactants, the term is used for linear molecules that are formed *via* imidazole-type intermediates (e.g. Figure 6).

Figure 6. An amphoteric surfactant (bottom) derived via *an imidazoline-type intermediate (right side on top).*



Most amphoteric imidazoline surfactants belong to chemical substance classes that are commonly known as amphodiacetates, amphodipropionates, or iminodipropionates. Therefore, these terms are used for the initial searches, yielding only very few results. On the other hand, search for the characteristic SMILES code fragment, CC(=O)NCCN, results in more than 150 substances. However, most of these belong to a different substance class (imino- or aminoethyleneamides), which can be described by a SMILES denotation containing the same characteristic fragment. Only a small number of obtained hits can be validated as typical imidazolines. By extracting additional possible name fragments from the initial results, such as "imidazolium compounds", "amino]propanoate" and "β-alaninate", the list of representative substances can be further increased.

1D.3 Modified fatty amino ethylates
The term "fatty amine ethylates" is synonymous with "fatty amine ethoxylates", which are covered under 1A.4. Further clarification is needed regarding the "modifications" that are supposed to be present in this context.

3.2 Auxiliaries and finishing agents for fibre and yarn manufacturing

2A. Lubricants
2A.1 Mineral oils
The term "mineral oil" describes mixtures of longer-chain alkanes obtained from a mineral source, typically from petroleum by distillation. A synonym is "paraffin oil". Simple searches performed with both trivial names "mineral oil" and "paraffin oil" give already sixty results. Therefore, no further searches need to be considered.

2A.2 Ester oils
The term "ester oil" refers to a class of synthetic oils that are chemically synthesized from carboxylic acids and alcohols. Various structural sub-groups can be distinguished (Fahl 2000):

- Monoesters, generated from a monoalcohol, often a fatty alcohol, and a monocarboxylic acid, often a fatty acid: These substances are covered under 7D.2.
- Dicarboxylic acid esters, generated from monoalcohols and dicarboxylic acids, e.g. hexanedioic acid: For these substances, a simple search is performed using the term "dicarboxylic acid", and the results are then filtered by auto filter and manually.
- Polyolesters, generated from a diol, e.g. neopentylglycol, or a triol, e.g. trimethylolpropane, and two or three monocarboxylic acids, often fatty acids: As a simple search for the term "polyol ester" gives only one hit, further searches are performed with the search terms "triester" and "tetraester." The results are filtered manually.

- Complex esters generated from combinations of multivalent alcohols and mono- and dicarboxylic acids: This structure is considered too complex to allow an application of the structure-based approach for the identification of characteristic substances.
- Carbonates, generated from carbonic acid and two monoalcohols: A simple search is performed on the name fragment “yl carbonate”, followed by further refinement using auto filter and manual selection.
- Aromatic esters, generated from aromatic (poly)carboxylic acids and aliphatic alcohols, often C8-C10 fatty alcohols: These substances are covered under 6F.4.

2A.3 Polyalphaolefins (PAO)
Polyalphaolefins are made by polymerizing alpha-olefins, i.e. molecules with a double bond between carbon atoms 1 and 2. The resulting polymers are branched, saturated hydrocarbons. A simple search using the general term “polyalphaolefin” gives only a small number of results. Due to the complexity of the polymeric structure no further search strategies can be devised.

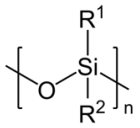
2A.4 Dialkyl benzenes
This substance class contains benzene derivatives with two (identical or different) alkyl substituents. A search for the simple term “alkyl benzene” generates only few results. On the other hand, the name fragment “ylbenzene” resulted in >3,600 hits. Excluding most additional functional groups, this number can be reduced to a couple of hundred and further refined by manual selection. An additional search based on the SMILES code fragment Cc1ccccc1 does not produce any additional new results. The retrieved representative substances are sorted according to the chain length of the alkyl substituents.

2A.5 Synthetic esters
This substance class is covered under 2A.2 Ester oils.

2A.6 Polyglycols
Polyglycols are formed by linking varying numbers of small diol units, mostly ethanediol or propanediol, via ether bonds. The terminal end can be a free hydroxy group or an ether with a different (longer chain) alcohol unit. After an unsuccessful initial search using “polyglycol”, results are obtained with the search term “glycol ether”. The SMILES code fragments OCCOCCOCCO (for ethylene-based ethers) and OCCCOCC (for both propylene and mixed propylene/ethylene ethers) give additional hits. The retrieved substances are sorted into two groups: ethylene-based substances and propylene-based substances.

2A.7 Silicones
Silicones (polysiloxanes) consist of chains of alternating silicon and oxygen atoms, in which the silicon atom bears two organic or halogen substituents (Figure 7).

Figure 7. Polysiloxane.



Simple searches with the terms “siloxane” and “silicone”, respectively, give several hundred results. Without further information regarding additional structural features that may be relevant in

this context, a further refinement does not seem possible. A small selection of substances is retrieved by a simple search using the term “silicone oil”.

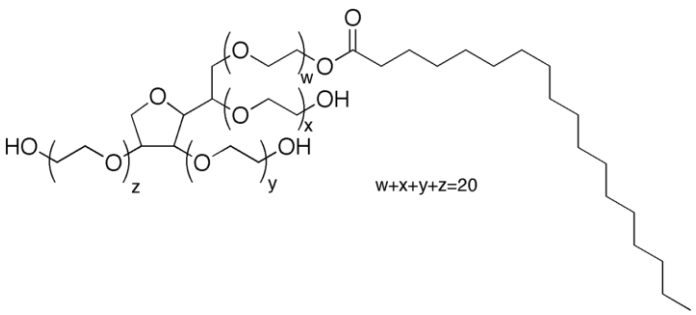
2A.8 EO/PO adducts
The term “EO/PO adducts” is synonymous with “ethylene oxide/propylene oxide adducts”, which are covered under 1A.6.

2B. Emulsifiers
2B.1 Ethoxylated fatty alcohols
The term “ethoxylated fatty alcohols” is synonymous with “fatty alcohol ethoxylates”, which are covered under 1A.1.

2B.2 Ethoxylated fatty acids
The term “ethoxylated fatty acids” is synonymous with “fatty acid ethoxylates”, which are covered under 1A.2.

2B.3 Ethoxylated sorbitan esters
These substances are derived from the heterocyclic polyol sorbitan, a condensation product of the sugar alcohol sorbitol. In ethoxylated sorbitan esters, the hydroxy groups are linked to (poly)ethylene ether groups of varying chain length and to one or more fatty acid moieties (Figure 8). A simple search with the general term “sorbitan” and manual processing of the initial results already gives a sufficient number of representative substances.

Figure 8. Polysorbate 60 (polyoxyethylene (20) sorbitan monostearate), an example for an ethoxylated sorbitan ester.



2B.4 Alkyl phenol ethoxylates (APEO)
Alkyl phenol ethoxylates (APEO) are already covered under 1A.3.

2B.5 Partial glycerides and ethoxylated triglycerides
Ethoxylated triglycerides are already covered under 1A.5. In order to identify partial glycerides (i.e. mono- and/or diglycerides), a search is performed with the term “glyceride”, yielding 638 results. These are refined by auto filter (contains “mono-“ or “di-“; does not contain “tri-“ and “ates”) and manual selection. The still considerable number of representative substances is sorted into groups depending on whether they are mono- or diglycerides or mixtures thereof and whether they are ethoxylated or not.

2B.6 Fatty amides
This substance class contains fatty acid derivatives of the generic formula $CH_3(CH_2)_nCONRR'$ with $n \geq 6$. Searches for the trivial name “fatty amide” or “fatty acid amide” give only a small number of

results. On the other hand, a search for the SMILES code fragment CCCCCCCC(=O)N produces almost a thousand results. These are further refined by adding "Sector of use = SU 5: Manufacture of textile, leather, fur" in the "Advanced search" function of the ECHA database.

2B.7 Sulphonated and sulphated vegetable oils
This substance class is already covered under 1B.9.

2C. Wetting agents
2C.1 Short-chain alkyl phosphates
This is a sub-group of the substance class "alkyl phosphates", which is covered under 2D.2.

2D. Antistatic agents
2D.1 Anionic surfactants
Anionic surfactants are discussed in detail in section 1B.

2D.2 Alkyl phosphates
Alkyl phosphates are phosphoric acid derivatives of the generic formula $(RO)_3PO$ that contain one to three organic residues. As this term is rather broad, the difficulty lies in defining an appropriate search strategy that does not result in too many hits. The search term "phosphate" relates to more than 3,600 substances in the ECHA database, which – in addition to organic phosphates – also include all kinds of inorganic salts. A search for "phosphoric acid ester", excluding salts, is more successful, but still gives more than a hundred valid results. The identified substances are sorted into groups according to the number of organic groups present in the molecule and their chain length.

2D.3 Sarcosides
These substances are derived from the amino acid sarcosine, also known as N-methyl glycine, by linking a fatty acid moiety to the nitrogen *via* an amide bond according to the generic formula $RCON(CH_3)CH_2COOH$, with R = fatty alkyl. A number of representative substances can be retrieved by a combination of searches based on the term "sarcosine", the name fragment „oylglycine", and the SMILES code fragment C(=O)N(C)CC(O)=O.

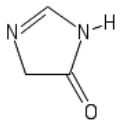
2D.4 Amine oxides
Amine oxides are oxidized tertiary amines of the generic formula $R_3N^+\text{-}O^-$. The three alkyl substituents can be identical or different. A simple search based on the term "amine oxide" already gives a sufficient number of representative substances, all of which belong to either of two groups: dimethylalkylamine oxides and bis(hydroxyethyl)alkylamine oxides.

2D.5 Sulpho succinates
These substances are covered under 1B.8.

2E. Biocidal additives
2E.1 Formaldehyde-containing compounds
The search term "formaldehyde" gives 1,448 results. In combination with the parameters "Regulatory context = Biocidal Product Regulation" and "Product category = PC8: Biocidal products", respectively, these can be narrowed down to less than ten.

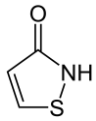
2E.2 Imidazolinones
These substances contain the heterocyclic core structure imidazolinone (Figure 9). A simple search for the generic name already results in a sufficient number of representative substances.

Figure 9. The imidazolinone structural core.



2E.3 Isothiazolinones
This widely used biocidal substance class is derived from the heterocyclic core structure isothiazolinone (Figure 10). The search term "isothiazolinone" gives only two results, but the more systematic "thiazol-3-one yields more than twenty substances.

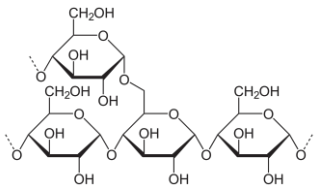
Figure 10. The isothiazolinone structural core.



3.3 Sizing agents

3A. Starch
Starch is a naturally occurring polysaccharide composed, to varying percentages, of the linear component amylose and the branched component amylopectin (Figure 11).

Figure 11. Amylopectin, a component of starch.



A simple search using the trivial name "starch" results in one hundred hits, which are narrowed down to seven by auto filter and manually selection.

3B. Starch derivatives
3B.1 Dextrins
Dextrins are smaller units derived from starch by partial hydrolysis. Simple search for the term "dextrin" results in more than eighty derivatives, which are processed by manual selection.

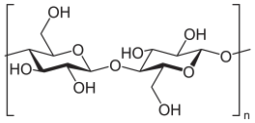
3B.2 Starch esters (most common: phosphates, acetates)
In these substances, the free hydroxy groups contained in starch form ester bonds with certain acids, mostly phosphoric acid and acetic acid, respectively. To identify representative substances,

the results initially retrieved under 3A are filtered for hits containing the terms “acet”, “phosph”, or “oate”.

3B.3 Starch ethers (most common: hydroxyethyl, hydroxypropyl, carboxymethyl)
In these substances, the free hydroxy groups contained in starch form ether bonds with certain short-chain alcohols. To identify representative substances, the results initially retrieved under 3A are filtered for hits containing the term “ether”.

3C. Cellulose derivatives (carboxymethyl cellulose)
Cellulose is a natural polysaccharide consisting of linear chains of linked glucose units (Figure 12). In cellulose derivatives, the free hydroxy groups can form ether or ester bonds with external components.

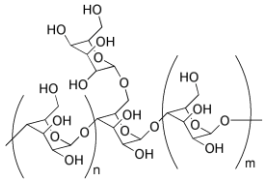
Figure 12. A segment of cellulose.



To identify representative substances, the results from a simple search using the term “cellulose” are filtered and sorted into three sub-groups: carboxy(methyl) derivatives, hydroxyalkyl derivatives, and others.

3D. Galactomannan derivatives (hydroxypropyl and carboxymethyl galactomannan)
Galactomannans are polysaccharides consisting of two different monomers: mannose units, linked in a linear chain, and galactose units, forming occasional sidechains.

Figure 13. A typical segment of galactomannan.



A search in the ECHA database using the term “galactomannan” does not give any valid results. Due to the complexity of the structure, no other search terms can be defined.

3E. Polyvinyl alcohol (PVA)
3E.1 Partially hydrolysed PVA (ca. 88%)
3E.2 Fully hydrolysed PVA (ca. 98%)
3E.3 Co-polymers of PVA with methyl methacrylate or other carboxyl monomers
Polyvinyl alcohol (PVA) is a synthetic polymer that is described by the generic formula $[CH_2CH(OH)]_n$. It is prepared by hydrolysis of polyvinyl acetate, $[CH_2CH(OCOCH_3)]_n$. In partially hydrolysed PVA, some of the hydroxy groups remain acetylated. In fully hydrolysed PVA, practically all hydroxy groups are free. In co-polymers of PVA with other carboxyl monomers, some of the hydroxy groups form ester bonds with carboxylic acids other than acetic acid. Simple searches performed on the terms “polyvinyl alcohol”, “polyvinylalcohol”, “PVA”, and “ethenol” give only a

small number of results, which are grouped according to their structure. An additional search for the SMILES code fragment CC(O)CC(O) does not give any valid results.

3F. Poly(meth)acrylates

3F.1 Considerable variation in chemical structure (building blocks: acrylic acid, acrylic esters, acrylamide, acrylonitrile, methacrylic acid)

Polyacrylates and polymethacrylates are polyesters made from derivatives of acrylic acid (propenoic acid) and methacrylic acid (2-methylpropenoic acid), respectively. A number of representative substances can be identified by simple searches for the terms “polyacrylate” and “polymethacrylate”, respectively.

3G. Polyesters

3G.1 Condensates of aromatic dicarboxylic acids with diols (e.g. ethylene glycol, diethylene glycol) and sulfonated aromatic dicarboxylic acids

Some of these substances can be identified by manual selection of the search results from a simple search with the term “polyester”.

3.4 Detergents/wetting agents

The substance classes under this heading are already covered elsewhere:

- 4A. Non-ionic surfactants: see 1A. Non-ionic surfactants
- 4B. Anionic surfactants: see 1B. Anionic surfactants

3.5 Auxiliaries containing sequestering agents

5.1 Ethylenediamine tetraacetate (EDTA)

5.2 Nitrilotriacetate (NTA)

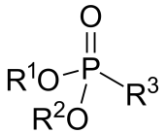
5.3 Diethylenetriamine pentaacetate (DTPA)

The above substances are listed individually, not as part of more broadly defined substance classes. Therefore, no further searches are warranted.

5.4 Phosphonic acid derivatives (phosphonates)

These can be salts or esters of phosphonic acids (Figure 14). This substance class is very large: simple searches with the terms “phosphonat” and “phosphonic acid” give 1,167 and 997 results, respectively. As no further information is available regarding the specific structure relevant for the textile industry, the results are narrowed down by applying the parameter “Sector of use =SU5: textile, leather, fur” in the advanced search function of the ECHA database. This way, over twenty characteristic substances can be identified.

Figure 14. Phosphonic acid derivatives.

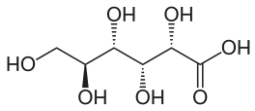


5.5 Gluconic acid derivatives (gluconates)

These can be salts or esters of gluconic acid (Figure 15), a natural compound derived from D-glucose. Simple searches are performed on the terms “gluconat” and “gluconic acid”. The results are

filtered manually and sorted into two groups: gluconic acid salts (approximately sixty compounds) and esters and other derivatives (eighteen compounds).

Figure 15. Gluconic acid.



5.6 Polyacrylates
These substances are covered under 3F.

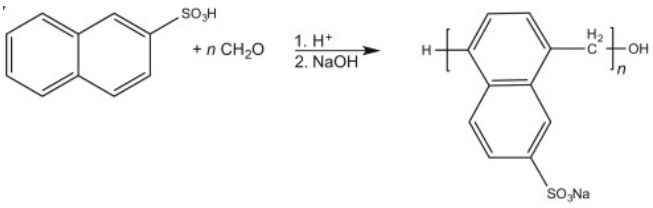
3.6 Dyeing auxiliaries

6A. Wetting, penetrating and de-aerating agents
The substance classes listed under this heading are already covered elsewhere:

- 6A.1 Alcohol polyglycol ethers: see 1A.1 Fatty alcohol ethoxylates
- 6A.2 Polyglycol esters: see 1A.2 Fatty acid ethoxylates
- 6A.3 Alkane sulphones: see 1B.1 Alkyl sulphonates
- 6A.4 Ethoxylated amines: see 1A.4 Fatty amine ethoxylates

6B. Dispersing agents
6B.1 Condensation products of naphthalene sulphonic acid with formaldehyde
Naphthalene sulphonic acid – or its derivatives – can be condensed with formaldehyde to give polymeric structures according to Figure 16. To identify these substances, the results from a simple search for "formaldehyde" are filtered for entries containing the name fragments "naphthalen" and "sulph" (or "sulf"). The retrieved substances are sorted into three groups: those containing unsubstituted naphthalene substructures, those containing alkylated naphthalene substructures, and those containing naphthalene substructures in addition to other structural components, e.g. cresol, hydroxybenzene or urea.

Figure 16. Condensation of naphthalene sulphonic acid with formaldehyde



The other substance classes listed under the heading "Dispersing agents" are already covered elsewhere:

- 6B.2 Lignosulphonates: see 1B.4 Ligninsulphonates
- 6B.3 Anionic surfactants: see 1B. Anionic surfactants
- 6B.4 Non-ionic surfactants: see 1A. Non-ionic surfactants

6C. Levelling agents
The following substance classes under this heading are already covered elsewhere:

- 6C.1 Fatty alcohol ethoxylates: see 1A.1

- 6C.2 Fatty amine ethoxylates: see 1A.4
- 6C.3 Fatty acid ethoxylates: see 1A.2
- 6C.4 Alkyl phenol ethoxylates (APEO): see 1A.3
- 6C.5 EO/PO adducts: see 1A.6 Ethylene oxide/propylene oxide adducts
- 6C.6 Fatty alcohol sulphates: see 1B.5 Alkyl sulphates

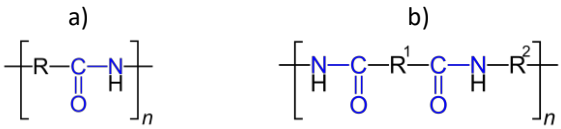
6C.7 Alkylaryl sulphates

This substance class contains esters of sulfuric acid with alkyl-substituted phenols, as in the general formula R-C_6H_4-O-SO_2-O^-, thus resembling the substance class of alkylaryl sulphonates that is covered in 1B.2. As evidenced by the unsuccessful initial search, these substances cannot be identified by the generic name as such. The name fragment “ylbenezenesulphate”, which was tried in analogy to the search performed for alkylaryl sulphonates before, remains equally ineffective. Instead, searches are performed on the combination of the terms “sulfate/sulphate”, in either spelling, and “phenyl”. The several hundred results can be effectively filtered, mostly by excluding halogen and/or nitrogen substituents, to obtain a small number of characteristic entries.

6C.8 Polyamide amines

Polyamides are polymers obtained by condensation of amino acids (Figure 17a) or of dicarboxylic acids with diamines (Figure 17b). In polyamide amines, the residues R, R^1 or R^2 contain additional free amino functions. A simple search for “polyamide amine” only leads to one entry: polyamide amine, CAS No. 67703-99-5, with no substance-specific information. Due to the complexity of the polymeric structure, no further search strategies can be devised.

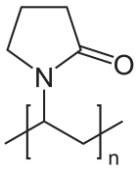
Figure 17. Polyamides formed from amino acids (a) and from dicarboxylic acids and diamines (b).



6C.9 Polyvinylpyrrolidone

These substances are linear polymers of vinylpyrrolidone (1-ethenyl-2-pyrrolidone) which differ only with regard to chain length (Figure 18). Simple searches can be performed using both the generic substance class name and the fragment “1-ethenyl-2-pyrrolidon”. Definition of a SMILES code fragment does not seem feasible.

Figure 18. Polyvinylpyrrolidone.



6C.10 Quaternary ammonium salts with C12-C14 fatty alkyl side chains

This is a subgroup of quaternary ammonium compounds, which are cationic surfactants (see 1C.1). As the name specifies that at least one of the four substituents should be C12-C14 alkyl, the several hundred results obtained from a simple search for “quaternary ammonium” can be effectively filtered using the descriptors “C12-C14” or “coco” (as coconut oil predominantly consists of C12-

and C14-saturated fatty acids). The results are sorted into subgroups according to whether or not they contain an aromatic ring system in addition to the fatty alkyl side chain. Those that do also fit into category 6C.11 (see below).

6C.11 Quaternary ammonium salts with aromatic ring systems
This is another subgroup of quaternary ammonium compounds. As the name specifies that at least one of the four substituents should contain an aromatic ring system, the results obtained for "quaternary ammonium" are filtered using the descriptors "phenyl" or "benzyl". The retrieved substances are sorted into subgroups according to the nature of the remaining three substituent. The majority of them fit into one of three structural categories: benzylalkyldimethyl, benzyldialkylmethyl, or benzylalkylbis(hydroxymethyl), with alkyl ranging from C7 to C22.

6C.12 Ethoxylated castor oil
Castor oil is a natural product obtained from castor beans. Its main component is the triglyceride of ricinoleic acid (12-hydroxy-9-*cis*-octadecenoic acid). A simple search for the trivial name results in more than 200 substances, which are manually filtered for all entries referring to ethoxylates, ethylene glycol esters or reaction products with ethylene oxide. Due to the complexity of the structure, no further search strategies are devised.

6C.13 Ethoxylated stearic acid
Stearic acid is the trivial name for octadecanoic acid. Searches are performed for both the trivial and the systematic name, yielding in each case around 350 hits. These are refined by Excel auto filter to retrieve entries containing the name fragments "ethoxyl" or "ethylene glycol" or "ethylene oxide" or "poly(oxy". This procedure leaves seven representative substances, which are considered sufficient as the definition of this substance class is rather narrow.

6C.14 Mixtures of alcohols, esters or ketones of medium chain length with emulsifying systems
A search is not possible, as this is not a structurally defined substance class, but an unspecified mixture of different substances.

<u>6D. Acid donors</u>
6D.1 Organic acid esters
For the identification of representative substances this term is too general. Further information would be needed regarding the structure of either the acid or the alcohol moiety of relevant substances. In part, this substance class is covered under 7D.2 Fatty acid esters.

6D.2 Fatty alcohol ethoxylates
This substance class is covered covered under 1A.1.

6D.3 Alkyl aryl sulphonates
This substance class is covered under 1B.2.

<u>6E. Antifoaming agents</u>
6E.1 Based on silicone derivatives
Silicone derivatives are covered under 2A.7.

6E.2 Tributylphosphates

This is a very specific subgroup of alkyl phosphates (see 2D.2). Only two substances can be found which match the substance class name: tributyl phosphate, $[CH_3(CH_2)_3O]_3PO$, and triisobutyl phosphate, $[CH_3CH(CH_3)CH_2O]_3PO$.

6E.3 Phosphoric esters
The term "phosphoric esters" is synonymous with "alkyl phosphates", which are covered under 2D.2.

6E.4 High molecular-weight alcohols
The name of this substance class is rather ambiguous, as it is not defined what should be considered "high molecular weight", if the alcohols should be primary, secondary or tertiary, or if the alkyl chain should be saturated or unsaturated, linear or branched. A chain length of eight carbon atoms is chosen as a cut-off value, as this corresponds to the most frequently used definition for "fatty alcohols". Searching for the trivial name "fatty alcohol" only gives a small number of characteristic substances. On the other hand, a search using the characteristic SMILES code fragment, CCCCCCCCO results in 1,500 hits, which are refined by filtering for the name fragment "nol". The retrieved entries contain mostly saturated linear alcohols of chain lengths between C8 and C28, as well as a few unsaturated representatives.

6E.5 Fluorine derivatives
For the identification of representative substances this term is too general. Further information would be needed regarding the structure of the fluorine-containing molecules.

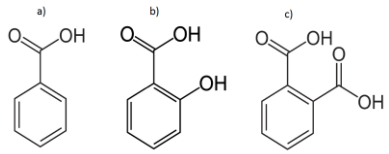
6F. Carriers
6F.1 Halogenated benzenes
Halogenated benzenes comprise a very large substance class. Even after excluding additional heteroatoms such as sulphur or nitrogen and functional groups such as carboxylic acids, around 150 representative substances are obtained from the simple searches for "chlorobenzene" and "bromobenzene" alone. The results are sorted into several groups depending on the type of halogen atoms and alkyl or alkoxy side chains they contain. In order to further narrow down this rather extensive list, more information would be needed regarding the nature of compounds that are commonly used as carriers for textile dyeing.

6F.2 Aromatic hydrocarbons (such as methylnaphthalene, diphenyl, trimethyl benzene)
As "aromatic hydrocarbons" in general would comprise a very large substance class, the searches are limited to the specific aromatic structures specified in brackets. Accordingly, simple searches are performed using the general term "aromatic hydrocarbon" as well as the terms "methylnaphthalene" and "trimethylbenzene". The several hundred results can be narrowed down to around 80 representative substances by automatic and manual filtration. A simple search for the term "diphenyl" results in more than 3,000 hits, which are not further processed, because it does not seem feasible without additional information.

6F.3 Phenols (such as o-phenylphenol, benzylphenol etc.)
A simple search for the name fragment "phenol" results in almost 6,000 matches. Therefore, the search is limited to the compounds specified in brackets. Searches for the terms "phenylphenol" and "benzylphenol" result in thirteen representative substances. Given the rather specific substructures, this number is considered adequate and no additional search strategies are devised.

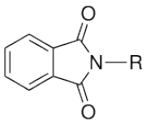
6F.4 Aromatic carboxylic acids and their esters (such as benzoates, salicylates, or phthalates)
Searches are limited to derivatives of the three aromatic carboxylic acids specified in brackets: benzoic acid, salicylic acid, and phthalic acid, respectively (Figure 19). Simple searches for the trivial names "benzoic acid" and "benzoate", respectively, result in several thousand matches. These are narrowed down to ten by adding the category "Sector of use = SU 5: Manufacture of textile, leather, fur" in the sub-menu "Uses and exposure" in the ECHA database "Advanced search" function, followed by manual sorting. The terms "salicylate" and "phthalate" respectively give a few hundred results, which are filtered using Excel auto filter and manual selection. In the case of salicylates, forty characteristic substances can be retrieved. However, in the case of phthalates almost two hundred substances remain, even after extensive selection, which cannot be narrowed down further without additional information regarding the specific type of phthalates used as dyeing carriers in the textile industry.

Figure 19. Benzoic acid (a), salicylic acid (b), and phthalic acid (c).



6F.5 Alkyl phthalimides (such as N-butylphthalimide)
Phthalimide is a trivial name for the bicyclic compound 2,3-dihydro-1H-isoindole-1,3-dione. In alkyl phthalimides, an alkyl chain is attached to the nitrogen atom (Figure 20). Simple searches are performed both with the systematic name and the trivial name, and in each case the results are further refined by auto filter and manual selection, leaving around twenty characteristic substances.

Figure 20. N-Alkylphthalimide.



3.7 Printing auxiliaries

7A. Thickening agents
7A.1 Natural polysaccharides, unmodified or modified
To some extent, this substance class is covered in sections 3A-3D. In order to perform additional searches, more information would be needed regarding the specific structures that may be relevant in this context.

7A.2 Synthetic polymers based on polyacrylic acid
This substance class is covered under 3F.

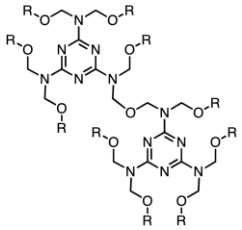
7A.3 Mineral oils
This substance class is covered under 2A.1.

7A.4 White spirit
"White spirit" is a trivial term for mixtures of medium-chain hydrocarbons obtained from petroleum or mineral oil. A simple search is performed on the term as such, and the results are listed as they are retrieved without further processing.

7B. Binders
7B.1 Polymer dispersions based on acrylates, butadiene, or vinyl acetate
Polyacrylates are covered in section 3F. Entries for polymers based on butadiene or vinyl acetate are retrieved by simple searches for the terms "vinyl acetate polymer", "poly vinyl acetate", and "butadiene acetate".

7C. Fixing agents
7C.1 Melamine-formaldehyde condensates
These are resins made from melamine (1,3,5-triazine-2,4,6-triamine) and formaldehyde (Figure 21). A simple search using the term "melamine formaldehyde" gives only a small number of results. But with the combination of the systematic name "1,3,5-triazine-2,4,6-triamine" and the name fragment "form", more than thirty characteristic polymers can be identified.

Figure 21. Sub-unit of a melamine-formaldehyde resin.



7D. Plasticisers
7D.1 Silicones
This substance class is covered under 2A.7.

7D.2 Fatty acid esters
This substance class contains esters of long-chain carboxylic acids with unspecified alcohols, e.g. $CH_3(CH_2)_nCOOR$ (but the acid may also contain double bonds). A simple search with the trivial name "fatty acid ester" gives only a small number of results. On the other hand, a search performed with the characteristic SMILES code fragment CCCCCCCC(=O)OC results in almost a thousand hits. To narrow these down, the category "Sector of use = SU 5: Manufacture of textile, leather, fur" is added in the sub-menu "Uses and exposure" in the ECHA database "Advanced search" function.

7E. Emulsifiers
7E.1 Aryl polyglycol ethers
In part, this substance class is covered under 1A.3 "Alkylphenol ethoxylates".

7E.2 Alkyl polyglycol ethers
The term "alkyl polyglycol ethers" is synonymous with "fatty alciohol ethoxylates", which are covered under 1A.1.

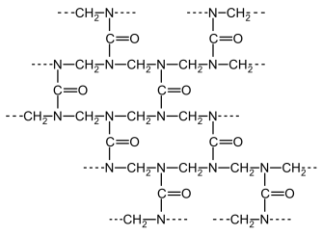
3.8 Finishing auxiliaries

8A. Cross-linking agents

8A.1 Agents based on formaldehyde and hydroxymethyl urea or methoxymethyl urea

Formaldehyde (HCHO) and urea (H_2NCONH_2) form polymers with the repeating unit $[(O)CNHCH_2NH]_n$ (Figure 22). A simple search using "urea formaldehyde" results only in a small number of hits, none of which are valid representatives of this substance class. Better results are obtained by combining a simple search for "urea" (resulting in 2,855 hits) with the Excel auto filter, filtering for entries that contain the fragments "formaldehyde" and "oxymethyl".

Figure 22. Segment of a cross-linked urea-formaldehyde polymer.



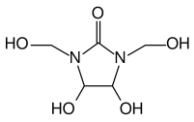
8A.2 Agents based on hydroxymethyl melamine or methoxymethyl melamine

In part, this substance class is covered under 7C.1 "melamine-formaldehyde condensates".

8A.3 Agents based on derivatives of hydroxymethyl dihydroxyethylene urea

Bis(hydroxymethyl)dihydroxyethylene urea (also known as dimethyloldihydroxyethylene urea or as 4,5-dihydroxy-1,3-bis(hydroxymethyl)-2-imidazolidinone) is a cyclic amide (Figure 23) that forms the basis of cross-linked polymer resins similar to those covered in 8A.1. It is not possible to identify any polymers belonging to this substance class. While the search terms "ethyleneurea" (and alternative spellings) and "imidazolidinone" lead to a few dozen results, none of these are polymers. Due to the complexity of the structure, a search based on SMILES code does not seem practical.

Figure 23. Bis(hydroxymethyl)dihydroxyethylene urea.



8A.4 Agents based on dimethyl urea and glyoxal

A similar search strategy as in 8A.1 can be employed: the results obtained for "urea" are filtered for those entries that contain one of the following terms: glyoxal, ethanedial, oxalaldehyde or oxaldehyde. After manual selection, only three valid results remain. An additional search based on the term "glyoxal" does not produce any additional hits.

8B. Catalysts

8B.1 Magnesium chloride

This substance is listed individually, no further search is warranted.

8B.2 Organic acids (e.g. alpha-hydroxycarboxylic acids)

The definition of this substance class is too general, a search is not possible.

8B.3 Inorganic Lewis acids
A search is not possible, because no indication is given regarding the composition and structure of these substances.

8B.4 Ammonium salts (e.g. chloride, sulphate and nitrate)
A search for "ammonium" results in almost 6,000 hits. It does not seem feasible to proceed without additional information.

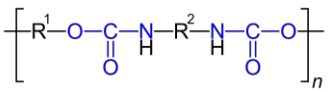
8C. Additives (softeners, stiffening agents, etc.)
8C.1 Polyacrylates
These substances are covered under 3F.

8C.2 Polysiloxanes
These substances are covered under 2A.7.

8C.3 Polyethylene waxes (partially oxidised polyethylenes)
Polyethylenes are synthetic polymers of the generic formula $(CH_2CH_2)_n$. In polyethylene waxes, the chain length n ranges from25 to 50. Oxidised polyethylenes contain polar features such as hydroxyl or carboxyl groups. Using the search terms "polyethylene wax" and "oxidised polyethylenes", a small number of entries can be identified.

8C.4 Polyurethanes
Polyurethanes contain small organic units such as diols and isocyanates which are linked through a urethane bond (Figure 24). A simple search for the term "polyurethane" is sufficient to identify almost twenty characteristic substances.

Figure 24. Linear polyurethane structure.



8C.5 Fatty acid condensation products (cationic softeners)
A search is not possible – more information is needed regarding the composition and structure of these substances.

8C.6 Polyvinyl acetate
Polyvinyl acetate is a polyester consisting of ethenyl acetate units. Searches for the general term "polyvinyl acetate" and for the more systematic fragment "ethenyl acetate" result in twelve characteristic substances.

8D. Biocides
8D.1 Permethrin
8D.2 Cyfluthrin
8D.3 Sulcofuron

The above substances are listed individually, not as part of more broadly defined substance classes. Therefore, no further searches are warranted.

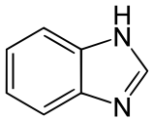
8D.4 Zinc organic compounds
These are compounds, in which zinc is linked to a carbon atom by a covalent or coordinate bond. The term "zinc organic compounds" does not refer to simple salts. Using the Advanced Search menu in the ECHA database, the name component "zinc" (which by itself corresponds to more than two thousand entries) was combined with either "Product Category 8: Biocides" or "Sector of Use 5: Manufacture of textiles, leather, fur". This way, a small number of representative substances are obtained.

8D.5 Tin organic compounds
These are organic compounds containing tin in a covalent or coordinate bond. Searches are carried out in the Advanced Search menu in analogy to 8D.4 above.

8D.6 Dichlorophenyl(ester) compounds
The name of this substance class relates to all molecules containing a structural unit derived from dichlorophenol, $C_6H_3Cl_2OR$. A simple search using the term "dichlorophenyl" gives more than two thousand results. These are refined through the Advanced Search function by adding "Product Category 8: Biocides", "Sector of Use 5: Manufacture of textiles, leather, fur" or "Regulatory context: BPR (Biocidal Products Regulation)". This way, a small number of representative substances can be retrieved.

8D.7 Benzimidazole derivatives
Benzimidazoles are heterocyclic aromatic compounds combining the structural features of benzene and imidazole (Figure 25). This substance class is quite large – a simple search using the term "benzimidazol" results in more than eight hundred matches. These are narrowed down to fifteen through the Advanced Search function by adding "Sector of Use 5: Manufacture of textiles, leather, fur" and "Regulatory context: BPR (Biocidal Products Regulation)", respectively.

Figure 25. Benzimidazole.



8D.8 Triclosan
This substance is listed individually, no further search is warranted.

8D.9 Isothiazolinones
This substance class is covered under 2E.3

8E. Antistatic agents
The substance classes under this heading are already covered elsewhere:

- 8E.1 Quaternary ammonium compounds: see 1C.1
- 8E.2 Alkyl phosphates: see 2D.2
- 8E.3 Alkyl ether phosphates: see 1B.11

8F. Flame retardants

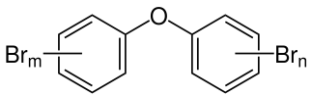
8F.1 Diammonium phosphate
8F.2 Aluminium sulphate
8F.3 Ammonium sulphate
8F.4 Aluminium hydroxide
The above substances are listed individually, not as part of more broadly defined substance classes. Therefore, no further searches are warranted.

8F.5 Titanium halogenated salts
The name relates to ionic compounds that contain both titanium and a halogen component (i.e. fluorine, chlorine, bromine or iodine). The search for these substances is rather straight forward by combining the name fragments “titan” and “chlor OR brom OR fluor”. Around twenty characteristic substances can be retrieved, most containing either fluorine or chlorine.

8F.6 Zirconium halogenated salts (e.g. potassium hexafluorozirconate)
The name relates to ionic compounds that contain both zirconium and a halogen component. The search is performed in analogy to 8F.5. Again, most of the retrieved substances contain either fluorine or chlorine.

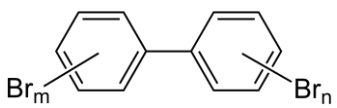
8F.7 Polybrominated diphenyl ethers (PBDE), e.g. pentabromodiphenyl ether (penta-BDE), octabromodiphenyl ether (octa-BDE), decabromodiphenyl ether (deca-BDE)
In this substance class, two brominated phenyl rings are linked through an ether bond (Figure 26). The number of bromine atoms is variable. A simple search for the term “bromodiphenyl ether” results in more than twenty characteristic compounds. As this is considered sufficient, no further search strategies are needed.

Figure 26. Polybrominated diphenyl ethers.



8F.8 Polybromo biphenyls (PBB), e.g. decabromobiphenyl
In this substance class, two brominated phenyl rings are linked directly (Figure 27). A simple search for the term “bromobiphenyl”, complemented with the alternative term “bromo-1,1’-biphenyl” and the name fragment “bromophenyl)benzene”, yields an overall number of twelve characteristic compounds.

Figure 27. Polybromobiphenyls.



8F.9 Tetrabromobisphenol A (TBBA)
8F.10 Hexabromocyclodecane
The substances are listed individually, no further search is warranted.

8F.11 Chloroparaffins

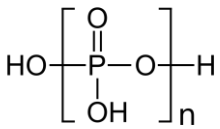
"Paraffin" is sometimes used as a synonym for "alkane", meaning a saturated hydrocarbon. Chloroparaffins contain a variable number of chorine aoms according to the general formula $C_xH_{(2x-y+2)}Cl_y$ The search terms "chloroparaffin" and "alkanes, chloro" only gives a small number of results. Searches for the general term "alkanes", followed by filtering the results for the name fragment "chloro", are more successful.

8F.12 Phosphate esters
Phosphate esters (alkyl phosphates) are already covered under 2D.2.

8F.13 Phosphine oxides
Phosphine oxides contain up to three organic residues covalently bound to a phosphorous atom, which is also linked to oxygen via a double bond, $R_3P=O$. The three residues can be identical or different. A simple search for the generic name "phosphine oxide" is sufficient to obtain around forty substances, which are grouped according to the structure of the organic substituents (i.e. aliphatic *versus* aromatic).

8F.14 Phosphoric acids
Phosphoric acid can exist in monomeric, dimeric, trimeric or even polymeric form (Figure 28). The monomer, H_3PO_4, is also known as orthophosphoric acid, the dimeric form as pyrophosphoric acid.

Figure 28. Phosphoric acids (n = 1: orthophosphoric acid; n = 2: pyrophosphoric acid).



8F.15 Phosphonic acid (2-((hydroxymethyl)carbamyl)ethyl)-dimethyl ester
This substance is listed individually, no further search is warranted.

8F.16 Phosphonium salt and urea precondensates
The description does not allow a systematic search. More information is needed in particular regarding the nature of the phosphonium compound.

8G. Repellents
8G.1 Wax-based repellents, typically containing paraffin and zirconium or aluminium salts
As the description relates to a mixture rather than a chemical substance class, a systematic search is not possible.

8G.2 Resin-based repellents (condensation products of fatty acids, alcohols or amines with methylolated melamines)
The description does not allow a systematic search.

8G.3 Silicone repellents, consisting of dimethylpolysiloxane, emulsifiers, glycols and water
As the description relates to a mixture rather than a chemical substance class, a systematic search is not possible.

8G.4 Fluorochemical repellents, mostly copolymers of fluoroalkyl acrylates and methacrylates
As the description relates to a mixture rather than a chemical substance class, a systematic search is not possible.

8H. Softeners
8H.1 Non-ionic surfactants
This substance class is covered under 1A.

8H.2 Quaternary ammonium compounds (such as stearyl or distearyl dimethyl ammonium chloride)
This substance class is covered under 1C.1.

8H.3 Amido amines formed by reaction of a fatty acid or a glyceride and a short chain polyamine
The description does not allow a systematic search.

8H.4 Imidazolines which can be acetylated or reacted with ethylene oxide
The description does not allow a systematic search.

8H.5 Polyethylene waxes
This substance class is covered under 8C.3.

8H.6 Silicones
This substance class is covered under 2A.7.

3.9 Coating compounds and auxiliaries

9A. Coating powders and pastes
9A.1 Polyolefins (e.g. polyethylene)
Part of this substance class is covered under 8C.3 (polyethylene waxes/oxidized polyethylenes). Additional substances can be identified with a simple search for the general term “polyolefin”.

9A.2 Polyamide 6
9A.3 Polyamide 6.6
These substances are listed individually, no further search is warranted.

9A.4 Copolyamides
The term copolyamides refers to polyamides that are generated from several different monomers. A direct search for the general term “copolyamide" does not give any results. A search for “polyamide” gives 58 hits. After manual selection, only two substances remain that are characteristic for this substances class.

9A.5 Polyester
These substances are covered under 3G.

9A.6 Polyurethane
These substances are covered under 8C.4.

9A.7 Polyvinylchloride
Even though there are different materials fitting the term “polyvinylchloride”, a simple search in the ECHA database gives only one characteristic entry.

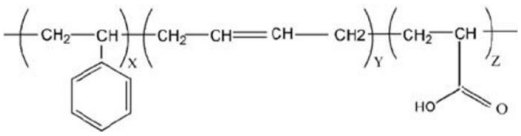
9A.8 Polytetrafluoroethylene
Listed as individual substance, no search warranted.

9A.9 Ethylene vinyl acetate
These are co-polymers consisting of ethylene and vinyl acetate units. Even though there are different materials fitting this description, a simple search in the ECHA database for “ethylene vinyl acetate” gives only one characteristic entry.

9A.10 Bitumen
This substance class comprises highly viscous fractions of petroleum, mostly consisting of aromatic hydrocarbons and heterocyclic compounds. Synonyms include asphalt, tar, and pitch. A simple search for the term “bitumen” gives 26 valid results. As this substance class contains heterogenous mixtures of substances, further structure-related searches are not possible.

9A.11 XSBR latex
XSBR is an acronym for carboxylated styrene-butadiene rubber, a polymer made from butadiene, styrene and an unsaturated carboxylic acid (Figure 29). Searches for the terms “XSBR” or “styrene butadiene” do not give any results. Therefore, the search results obtained for “styrene” and “butadiene”, respectively, are processed by Excel auto filter and manual selection. This way, seventeen characteristic substances can be identified.

Figure 29. Chemical structure of XSBR.



9B. Polymer dispersions (aqueous formulations)

9B.1 Poly(meth)acrylate
9B.2 Polyacrylic acid
9B.3 Polyacrylonitrile
9B.4 Polyacryloamide
These substance classes are covered in section 3F.

9B.5 1,3-Polybutadiene
Representatives for this substance class can be identified by a simple search for “polybutadiene” and by a search for “butadiene” followed by auto filter and manual selection.

9B.6 Polystyrene
Representatives for this substance class can be identified by a simple search for “polystyrene” and by a search for “styrene” followed by auto filter and manual selection.

9B.7 Polyurethane
This substance class is covered under 8C.4.

9B.8 Polyvinylchloride
This substance class is covered under 9A.7.

9B.9 Polyvinylacetate
This substance class is covered under 8C.6.

9B.10 Copolymers of the above-mentioned polymers
These substances are included in the respective sections above.

9C. Melamine resins produced by reaction of melamine and formaldehyde
These substances are covered under 7C.1.

9D. Polymer dispersions (organic solvent-based formulations)
9D.1 Polyurethanes
This substance class is covered under 8C.4.

9D.2 Silicones
This substance class is covered under 2A.7.

3.10 Solvents
This section contains substances which are listed individually, not as part of more broadly defined substance classes. Therefore, no further searches are warranted.

3.11 Other chemicals
This section contains substances which are listed individually, not as part of more broadly defined substance classes. Therefore, no further searches are warranted.

4. Extending the approach to other industry sectors

The identification of relevant substances used in the textile industry could be based on an extensive list of chemical substance classes typically associated with the technical processes in this sector. For other industries, this kind of structured information is not available. A number of case studies were carried out as part of the HAZBREF project in order to identify at least some chemicals used in different industry sectors. The objective of this study was then to define chemical substance classes, starting from those individual chemicals, and to apply the developed structure-based approach for the identification of wider selections of chemicals in each substance class that may be associated with certain industry operations.

The following case studies were available:

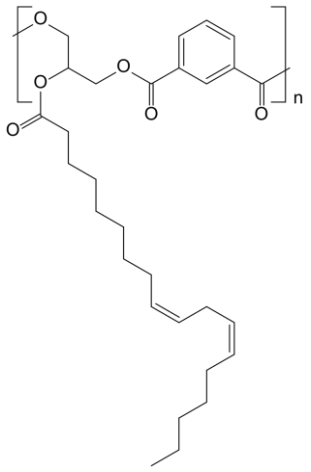
- Production of alkyd resins (HAZBREF 2020a);
- Production of polyolefins (HAZBREF 2019a);
- Surface treatment of metals (HAZBREF 2020b-e));
- Manufacture of fertilizers (HAZBREF 2019b).

It is to be noted that these case study reports only list chemicals that have already been identified as hazardous during classification and labelling or have been identified as substances of very high concern under REACH or as priority substances under the Water Framework Directive.

4.1 Production of alkyd resins

Alkyds are polyesters derived from polyols and di- or multivalent carboxylic acids (Figure 30). In addition, they often contain fatty acids linked to the polyol as a sidechain. If the fatty acid contains double bonds, cross-linking with modifiers such as styrene or acrylic acid derivatives is possible. Other modifiers such as epoxides, isocyanates, or silicones may be linked to hydroxy or carboxyl groups, either as sidechains or as insertions into the polymer backbone.

Figure 30. An example for an alkyd structure (composed of glycerol, isophthalic acid, and linoleic acid).



All chemicals mentioned in the available case study are summarized in Table 2, grouped according to their technical functions. They can be categorized as

a. building blocks that determine the structure of the produced polymer;
b. additives that determine some properties of the produced polymer;
c. process and maintenance chemicals.

Table 2. Substances used in the production of alkyd resins (case study from Akzo Nobel, Estonia).

Process/application	Name
Alkyd synthesis – carboxylic acid component	Maleic anhydride
	Phthalic anhydride
	Hexahydrophthalic anhydride
	Trimellitic anhydride
Alkyd synthesis – polyol component	Trimethylolpropane (propylidynetrimethanol)
	Glycerol (propane-1,2,3-triol)
	Pentaerythritol [2,2-bis(hydroxymethyl)propane-1,3-diol]
Alkyd synthesis – oil component	Conjugated fatty acids from sunflower oil (Nouracid SE30)
	Tall oil fatty acids (Sylfat 2)
	Distilled tall oils (Sylvatal 20S)
	Linseed oil
	Soyabean oil
Modifier	2,4-Toluene diisocyanate
	Acrylates
Biocide (preservative)	1,2-Benzisothiazol-3(2H)-one
	Tetramethylol acetylenediurea (1,3,4,6-Tetrakis(hydroxymethyl)tetrahydroimidazo[4,5-d]imidazole-2,5(1H,3H)-dione)
	Formaldehyde
	Benzoic acid (?)
Catalyst	Dibutyltin oxide
Polymerisation inhibitor	Hydroquinone
Neutraliser	Triethylamine
Surfactant (emulsifier)	ATSURF 3969 (polymeric emulsifier)
	Non-ionic polymeric surfactant (Maxemul 7101)
	Isopropylammonium C10-13 alkylbenzenesulfonate (Maxemul 7201-LQ)
	Sodium 4-diisotridecyl sulphonatosuccinate (AEROSOL TR-70 E)
Solvent	1-Butanol
	1-Butoxy-2-propanol
	Butyl acetate
	Ligroin
	White spirit
	Hydrocarbons C9-C11 (n-alkanes, isoalkanes, cyclics)
	Hydrocarbons C10-C13 (n-alkanes, isoalkanes, cyclics)
	Xylenes
	Limonene (1-methyl-4-prop-1-en-2-yl-cyclohexene)
	Diterpenes
Unknown function	Anthraquinone
	Benzophenone
	Benzoic acid (?)

In the case of polymer building blocks, a precisely defined chemical structure is critical for the technical function. It should therefore be possible to apply the previously developed structure-based approach for the identification of potentially relevant substances. For chemicals fulfilling

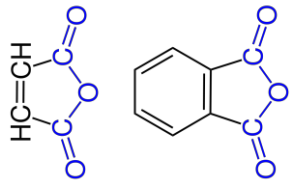
other technical functions, the application of the structure-based approached seems more limited, as will be discussed in the respective sub-sections below.

4.1A Alkyd polymer building blocks

A.1 Dicarboxylic acid anhydrides

As dicarboxylic acids in their native state are not sufficiently reactive to achieve condensation with the polyol hydroxy groups, alkyd production is based on the use of acid anhydrides. These are condensation products of two carboxylic acid units, under the elimination of water (Figure 31).

Figure 31. Maleic anhydride (left) and phthalic anhydride (right).



Simple searches were carried out with the terms “dicarboxylic acid anhydride” and “dicarboxylic anhydride”, resulting in a list of thirty representative substances. Most of them are derivatives of phthalic anhydride. In principle, all these substances would be suitable to function as polymer building blocks.

A.2 Polyols

Polyols are compounds that contain multiple hydroxy groups. As an initial search with the term “polyol” did not give any valid results, additional searches were performed using the name fragments “anetriol”, “-triol”, and “-tetrol”, respectively. After filtering the several hundred results, a list of about twenty triols and ten tetrols could be compiled, all which are in principle suitable to function as alkyd resin building blocks.

A.3 Fatty acids

Fatty acids are aliphatic carboxylic acids of the chain length C8-C24. They can contain alkyl side-chains and double bonds. Searches were performed on the trivial name “fatty acid” as well as on the name fragment “anoic acid”. In both cases, several thousand hits were obtained, which had to be filtered extensively. Still, almost three hundred substances remain, which in principle are representative of this substance class. In order to further narrow down these search results, additional industry-specific information would be required.

A.4 Vegetable oils

Vegetable oils are natural components of seeds consisting of fatty acids linked to a triglyceride core. A simple search for “vegetable oil” and “vegetable-oil”, respectively, gave around forty characteristic substances, which were sorted into the sub-groups “fats and oils”, “glycerides”, “Fatty acids”, “fatty acid salts”, and “fatty acid esters”.

A.5 Diisocyanates

These substances contain two isocyanate groups linked to an organic core structure, according to the generic formula O=C=N-R-N=C=O. They react with free hydroxy groups under the formation of amides or urethanes. A simple search for “diisocyanat” gave almost one thousand hits, which

could be narrowed down to around one hundred. The results were divided into two sub-groups: aliphatic and aromatic substances.

A.6 Acrylates
Acrylates are derivatives of acrylic acid (2-propenoic acid). A simple search for the common name gave more than two thousand results. It is not possible to narrow these down without additional sector-specific information.

4.1B Additives
Plastic products contain several auxiliary substances, called additives, which give them certain properties. Examples include stabilizers, plasticizers, flame retardants, or pigments. In a recently completed joint project, ECHA and industry collected information on the various known substances used as plastic additives for different purposes (https://echa.europa.eu/de/plastic-additives-initiative). The only kind of additives listed in the Akzo Nobel case study are preservatives, which are not covered by the above-mentioned study. However, as preservatives fall under the Biocidal Products Regulation and require authorisation before being placed on the market, they can be effectively identified in the ECHA database of biocidal substances. A structure-based search does not seem to be the best approach.

4.1C Process and maintenance chemicals
Some chemicals are used in the production process for specific functions. For instance, the alkyd resin case study lists the following:

- dibutyltin oxide as catalyst
- hydroquinone as polymerisation inhibitor
- triethylamine as neutralizer

It appears as if the listed chemicals are very specific for the respective purpose and not much variation by the use of structurally related substances should be expected. Therefore, no structure-based searches were performed on these substance classes.

Other auxiliary chemicals used in the production process fulfil rather generic functions, which are common to several different industry operations. In the alkyd resin case study report these are:

- *Surfactants/emulsifiers*
  These substances have been covered extensively in the previous chapters (see 3.1 Surfactants and 3.2B Emulsifiers).
- *Solvents*
  Solvents are practically used in all industry sectors, both during the production process and for the cleaning of facilities. According to the specific task, different types of solvents are employed. The solvents listed in the alkyd resin case study can be categorized as follows:
    - Hydrophilic: short-chain alcohols (1-butanol, 1-butoxy-2-propanol), butyl acetate
    - Hydrophobic: alkanes, terpenes, xylene

4.2 Production of polyolefins

Polyolefins are made by polymerisation of simple olefins (a synonym alkenes), such as ethene (polyethylene), or propene (polypropylene). Their production is covered in the adopted BREF for Polymers (European Commission 2007), in contrast to the alkyd resins discussed above, which are not included in this document. However, the Polymer BREF does not list a systematic collection of

chemical classes or individual chemicals used in polymer production. Table 3 summarizes the chemicals listed in the available case study (HAZBREF 2019a). The hydrocarbon co-monomers are circulated in the gas phase, and the auxiliaries such as catalysts and absorbents are bound to the solid phase. Therefore, all these components can be assumed to be of little, in any, relevance for industrial wastewater. Listed additives include primarily antistatics and maintenance chemicals. Antistatics are also covered in the European plastic additive initiative (see above). However, there is hardly any consistency between the list produced in that initiative and the one reported in this case study. Based on the available information, it was not possible to define structure-based chemical classes and perform searches for the identification of additional relevant chemicals for this industry sector.

Table 3. Substances used in the production of polyethylene and polypropylene

Process/application	Name
Co-monomer	Ethylene
	Propane
	Butene
	1-Hexene
Catalyst	Copper oxide
	Zinc oxide
	Aluminium
Absorbent	Copper oxide
	Zinc oxide
Additives:	
Antistatic	Methanol
	Toluene
	1,2,4-Trimethylbenzene
	Naphthalene
	Naphtha
	Heptane
	2-Methylhexane
	3-Methylhexane
	Methylcyclohexane
	3-Ethylpentane
	2,3-Dimethylpentane
	1-Decene
	Benzenesulfonic acid
	Dinonylnaphthalenesulphonic acid
	Quaternary ammonium compounds
	Dicocoalkyldimethylammonium salts
Maintenance chemical	Benzene
	Hydrocarbons C7
	Dimeric tall oil fatty acid
	Nonylphenol
	Triethylenetetramine
	2-Piperazinylethylamine
	Titanium dioxide
	Copper oxide
	Zinc oxide

4.3 Surface treatment of metals

This industry sector is covered by the adopted BREF for the Surface Treatment of Metals and Plastics (European Commission 2006). Most chemicals used in the surface treatment of metals (STM) are inorganic substances, such as metals and metal salts, inorganic acids, bases, and salts. To this kind of substances, the developed approach of defining chemical substance classes as a basis for the identification of structurally related chemicals is not very well applicable. Therefore, the following section will only consider organic chemicals reported from the STM industry, which are mainly used for purposes such as degreasing, cleaning, paint removal, lowering of surface tension, and sealing. The organic chemicals listed in the available case study reports are summarized in Table 4 to Table 7.

Table 4. Organic chemicals used in the surface treatment of metals (STM), Aurajoki case study

Process/application	Name
Biocide (preservative)	2-Methyl-2H-isotiazol-3-one
	5-Chlor-2-methyl-2H-isotiazol-3-one
Nickel coating	C12-14 alkylether sulfate with ethylene oxide, sodium salt
	3-(Amidinothio)propionic acid
	Sodium propanesulfonate
	Prop-2-yne-1-ol
Nickel plating	Prop-2-yne-1-ol, polymer with ethylene oxide
	But-2-yne-1,4-diol
	3-(Amidinothio)propionic acid
	Prop-2-yne-1-ol
Zinc coating	1-Methoxy-2-propanol
	4-Phenyl-3-butene-2-one
	Sodium p-cumenesulphonate
	Sodium benzoate
	1'-Acetonaphthone
	2-Butoxyethanol
	Alcohols, C9-11-iso-, C10-rich, ethoxylated
Aluminium washing	Amines, coco alkyl, ethoxylated
	Alcohols, C11-15-secondary, ethoxylated
	(Z)-9-Octadecenoic acid, sulfonated, potassium salts
Additive for alkaline zinc	1H-Imidazole, polymer with (chloromethyl)oxirane
	Bis(2-chloroethyl) ether-1,3-bis[3-(dimethylamino)propyl]urea copolymer
Tin plating	Polyethylene glycol mono(octylphenyl) ether
	1-Naphthaldehyde
Paint	Cyclohexane
	Butane
	Propane
	Xylene
	Ethylbenzene
Micro corrosion	p-Toluenesulfonic acid, including 5% sulphur acid
Surface removal	5-Chloro-2-methyl-2H-isothiazol-3-one/2-methyl-2H-isothiazol-3-one (3:1)
Cleaning	Amines, tallow alkyl, ethoxylated
	(Z)- 9-Octadecenylamine, ethoxylated

Process/application	Name
	(Z)- 9-Octadecenylamine, ethoxylated
Detergent	Isotridecanol, ethoxylated
	Fatty acids, C6-12, potassium salts
	Benzenesulfonic acid, C10-13-alkyl derivs., compds. with triethanolamine
	Amides, coco, N,N-bis(hydroxyethyl)
Maintenance	Hydrocarbons, C6, isoalkanes, <5% hexane
	Hydrocarbons C6-C7, n-alkanes, isoalkanes, cyclic
	Hydrocarbons C7, n-alkanes, isoalkanes, cyclic
	Carbohydrates, C7-C9, n-alkanes, isoalkanes, cyclic
	2-Propanol
unknown	Alkylether phosphate, 1 OE, sodium salt

Table 5. Organic chemicals used in the surface treatment of metals (STM), Estonian case study 1

Process/application	Name
Oiling of metal details	Hydrocarbons, C9-C11, n-alkanes, isoalkanes, cyclics, < 2% aromatics
	2-Propanol
Supress Cr vaporisation	Polyfluorosulphonic acid (3,3,4,4,5,5,6,6,7,7,8,8,8-tridecafluorooctanesulphonic acid)
Zincing, additive	4-Hydroxy-4-methylpentan-2-one
	4-Phenylbutenone
	2-(1-Methylethyl)benzenesulphonic acid, potassium salt (potassium cumenesulphonate)
	2-(1-Methylethyl)benzenesulphonic acid, sodium salt (sodium cumenesulphonate)
	Isodecanol, ethoxylated
	Sodium benzoate
	1-Acetonapthone
Electrochemical cleaning, additive	Alcohols C9-11, ethoxylated
	Hexyl diglucoside
	1,4-Butynediol
Wetting agent	Sodium 1,4-bis(1,3-dimethylbutyl)sulphonatosuccinate
	Sodium 2-ethylhexyl sulfate
	3-Hexyne-2,5-diol
	1,4-Butynediol
	2-Butyne-1,4-diol polymer with methyloxirane
	2-Propyne-1-ol polymer with ethylene oxide
	2,2,2-Trichloroethane-1,1-diol
	2-Propyne-1-ol
	3-Alkyl-C12-16-dimethylbenzyl-ammonium chloride

Table 6. Organic chemicals used in the surface treatment of metals (STM), Estonian case study 2

Process/application	Name
Steel cleaning	2-butoxyethanol
	2-(2-butoxyetoxy)ethanol
	Sulphonic acids, C14-16-alkane hydroxy and C14-16-alkene, sodium salts

Process/application	Name
Tin plating	1-Naphtaldehyde
	2-(2-[4-(1,1,3,3-Tetramethyl-butyl)phenoxy]ethoxy)ethanol
Unknown	1-Methoxy-2-propanol
	4-Phenyl-3-butene-2-on

Table 7. Organic chemicals used in the surface treatment of metals (STM), Polish case study

Process/application	Name
Anode sealing	Formaldehyde, reaction products with sulfonated 1,1'-oxybis[methylbenzene], sodium salts
Aluminium etching	Polyoxyethylene isodecyl ether
	2-Mercaptobenzothiazole
Paint removal	Benzyl alcohol
	Benzyl formate
	Mercaptobenzothiazole, sodium salt
	2-(heptadecenyl)-4,5-dihydro-1H-imidazole-1-ethanol
	Benzene, C10-C16 alkyl derivatives
Degreasing	Diethylene glycol butyl ether
	2-Mercaptobenzothiazole
	Alcohols, C12-15, ethoxylated
	Alcohols, C12-14, ethoxylated, propoxylated
	2,2'-(octadec-9-enylimino)bisethanol
Degreasing (vapour phase)	Tetrachloroethylene
	tert-Butyl glycidyl ether ((tert-butoxymethyl)oxirane)
	Dipropylene glycol N-butyl ether
	N-Methylmorpholine
Wax removal	Isotridecanol,ethoxylated
	Fatty acids, tall-oil (C18-C20), triethanolamine salts
	Butylhydroxytoluene
Emulsifier	Alcohols, C6-18, ethoxylated
Surface preparation/ nickel removal	3-Nitrobenzoic acid
	Ethylenediamine
	Sodium 3-nitrobenzenesulfonate
Chrome plating	3,3,4,4,5,5,6,6,7,7,8,8,8-tridecafluorooctanesulphonic acid
Cadmium plating	Sodium Gluconate
	Dextrin
	Polyvinyl Alcohol
Cleaning	Isoparaffinic hydrocarbon
Conservation	2-butoxyethanol
	Hydrocarbons C9-C11 (n-alkanes, isoalkanes, cyclics)
	Hydrocarbon waxes (petroleum), oxidized, methyl esters, barium salts
	4-Hydroxy-4-methyl-2-pentanone
	Benzenesulfonic acid, 2,3(or 3,4)-dimethyl-, mono-C10-14-branched alkyl derivs., sodium salts
	2-(2-Hepta-8-decenyl-2-imidazolin-1-yl)ethanol
Neutralization	Hydrotreated light distillate (petroleum)
	Diphenyl Isodecyl Phosphate
Penetration control	C11-15-Secondary alcohols, ethoxylated
	Terpineol
Dehydrogenation	Lubricating oils (petroleum)

Process/application	Name
Galvanizing	Triethanolamine
	Diethylenetriamine
	Diethanolamine
	N,N,N',N'-Tetrakis(2-Hydroxypropyl)ethylenediamine
Magnetic control	Distillates (petroleum), hydrotreated light
Unclear	Toluene
	Pentaerythritol
	Castor oil

It can be noted that some chemicals are listed in more than one case study and seem to be typical for this specific industry. The descriptions for the technical processes, in which the chemicals are applied are quite different between the different studies. Therefore, grouping of chemicals was performed based on their structure only, ignoring the technical functions for the time being. This issue may need to be revisited later.

Table 8 lists chemicals belonging to substance classes that are already covered in previous chapters. Chemicals listed in Table 9 have not been previously covered. Some of them can be combined into groups of structurally related chemicals, such as alcohols, ketones, or hydrocarbons. However, since the description of technical functions is so heterogenous within the four case studies, no judgement can be made at the moment how far the chemical structure is related to the technical function. This would be important information for a more refined definition of the respective structure-based chemical classes.

Table 8. Organic chemicals listed in STM case studies that correspond to previously covered substance classes.

Substance class	Reference to textile chemicals
Non-ionic surfactants	
Alcohols C9-11, ethoxylated	*Fatty alcohol ethoxylates (1A.1)*
Alcohols, C11-15-secondary, ethoxylated	
Alcohols, C12-14, ethoxylated, propoxylated	
Alcohols, C12-15, ethoxylated	
Alcohols, C6-18, ethoxylated	
Alcohols, C9-11-iso-, C10-rich, ethoxylated	
Polyoxyethylene isodecyl ether	
Isodecanol, ethoxylated	
Isotridecanol, ethoxylated	
(Z)- 9-Octadecenylamine, ethoxylated	*Fatty amine ethoxylates (1A.4)*
Amines, coco alkyl, ethoxylated	
Amines, tallow alkyl, ethoxylated	
2,2'-(octadec-9-enylimino)bisethanol	
Anionic surfactants	
Sulphonic acids, C14-16-alkane hydroxy and C14-16-alkene, sodium salts	Alkyl sulphonates (1B.1)
Sodium propanesulfonate	
Sodium p-cumenesulphonate	Alkylaryl sulphonates (1B.2)
2-(1-Methylethyl)benzenesulphonic acid, sodium or potassium salt	

Substance class	Reference to textile chemicals
Benzenesulfonic acid, 2,3(or 3,4)-dimethyl-, mono-C10-14-branched alkyl derivs., sodium salts	
Benzenesulfonic acid, C10-13-alkyl derivs., compds. with triethanolamine	
p-Toluenesulfonic acid, including 5% sulphur acid	
C12-14 alkylether sulfate with ethylene oxide, sodium salt	Alkyl sulphates (1B.5)
Sodium 2-ethylhexyl sulfate	
Sodium 1,4-bis(1,3-dimethylbutyl)sulphonatosuccinate	Dialkylsulphosuccinates (1B.8)
Fatty acids, C6-12, potassium salts	Alkyl carboxylates (1B.10)
Fatty acids, tall-oil (C18-C20), triethanolamine salts	
Alkylether phosphate, 1 OE, sodium salt	Alkylether phosphates (1B.11)
Emulsifiers	
Amides, coco, N,N-bis(hydroxyethyl)	Fatty amides (2B.6)
Diphenyl Isodecyl Phosphate	Alkyl phosphates (2D.2)

Table 9. Organic chemicals listed in STM case studies that are not previously covered.

Provisional chemical class
Alcohols:
2-Propanol
1-Methoxy-2-propanol
2-Propyne-1-ol
Benzyl alcohol
1,4-Butynediol
3-Hexyne -2,5-diol
2,2,2-Trichloroethane-1,1-diol
Pentaerythritol
Alkylethers:
2-Butoxyethanol
2-(2-butoxyetoxy)ethanol
Diethylene glycol butyl ether
Dipropylene glycol N-butyl ether
tert-Butyl glycidyl ether ((tert-butoxymethyl)oxirane)
Polyethylene glycol mono(octylphenyl) ether
2-(2-[4-(1,1,3,3-Tetramethyl-butyl)phenoxy]ethoxy)ethanol
2-Butyne-1,4-diol polymer with methyloxirane
2-Propyne-1-ol polymer with ethylene oxide
Polyvinyl Alcohol
Ketones:
4-Hydroxy-4-methyl-2-pentanone
4-Phenyl-3-butene-2-one
1'-Acetonaphthone
Amines:
Ethylenediamine
Diethylenetriamine
Diethanolamine

Provisional chemical class

Triethanolamine
N,N,N',N'-Tetrakis(2-Hydroxypropyl)ethylenediamine

Aliphatic hydrocarbons:

Propane
Butane
Cyclohexane
Hydrocarbons C6-C7, n-alkanes, isoalkanes, cyclic
Hydrocarbons, C7-C9, n-alkanes, isoalkanes, cyclic
Hydrocarbons C7, n-alkanes, isoalkanes, cyclic
Hydrocarbons C9-C11 (n-alkanes, isoalkanes, cyclics)
Hydrocarbons, C6, isoalkanes, <5% hexane
Hydrocarbons, C9-C11, n-alkanes, isoalkanes, cyclics, < 2% aromatics
Distillates (petroleum), hydrotreated light
Hydrotreated light distillate (petroleum)
Lubricating oils (petroleum)
Isoparaffinic hydrocarbon
Hydrocarbon waxes (petroleum), oxidized, methyl esters, barium salts

Aromatic hydrocarbons:

Xylene
Toluene
Ethylbenzene
Benzene, C10-C16 alkyl derivatives

Other:

(Z)-9-Octadecenoic acid, sulfonated, potassium salts
1H-Imidazole, polymer with (chloromethyl)oxirane
2-(2-Hepta-8-decenyl-2-imidazolin-1-yl)ethanol
2-Mercaptobenzothiazole
2-Methyl-2H-isotiazol-3-one
5-Chlor-2-methyl-2H-isotiazol-3-one
3-(Amidinothio)propionic acid
3-Alkyl-C12-16-dimethylbenzyl-ammonium chloride
3-Nitrobenzoic acid
Polyfluorosulphonic acid (3,3,4,4,5,5,6,6,7,7,8,8,8-tridecafluorooctanesulphonic acid)
Benzyl formate
Bis(2-chloroethyl) ether-1,3-bis[3-(dimethylamino)propyl]urea copolymer
Butylhydroxytoluene
Castor oil
Dextrin
Formaldehyde, reaction products with sulfonated 1,1'-oxybis[methylbenzene], sodium salts
Hexyl diglucoside
N-Methylmorpholine
1-Naphtaldehyde
Sodium 3-nitrobenzenesulfonate
Sodium benzoate
Sodium Gluconate
Terpineol
Tetrachloroethylene

4.4 Manufacture of fertilizers

This industry sector is based on the use of inorganic chemicals. An application of the structure-based approach was not possible.

5. Summary and conclusions

The objective of this study was to identify chemicals that are potentially used in the textile industry based on their chemical structure and starting from a list of substances classes compiled by industry. Of the 204 entries contained in this list, 37 refer to individual substances and thus are not considered any further. Another 63 entries turned out to be duplicates, sometimes listed under the same name, sometimes under a synonym, because substances such as emulsifiers, detergents, etc. are used for more than one technical function. Of the remaining 104 unique substance classes, 14 proved unsuitable for the application of structure-based search strategies. In some cases, these entries relate to mixtures of chemicals rather than one specific chemical class; in other cases, the given definition is too general, e.g. “fluorine derivatives” or “phosphonium salts”, and more information regarding the structural skeleton would be needed to develop a search strategy. However, for 90 substance classes, structure-related search strategies could be employed successfully to identify characteristic chemicals.

The most frequently applied search strategy was a simple search for the trivial or common name of the substance class. In 37 cases, no additional searches were performed. Only in nine case, the search for the common name failed to generate any valid results. In 13 cases, this approach did not seem feasible, because there are no known common names for certain substance classes. The second most successful search strategy was a simple search for a systematic name fragment applying to the substance class. This was often performed in addition to a trivial name search but could sometimes also serve as the only search strategy applied. In about half of the cases, a systematic name fragment was not employed – either because the trivial name search had already generated a sufficient number of results or because it was not possible to derive a systematic name fragment for certain substance classes. This is particularly true for complex natural products such as lignosulphonates or polysaccharides, for mixtures of petroleum-derived substances such as white spirit or bitumen, and for some synthetic polymers. An advanced search for a SMILES code fragment was successfully applied to twenty substance classes. These were mostly linear molecules with an easily denoted central functional group, such as alcohol ethoxylates, carboxylates, alkyl phosphates, alkyl sulphates or sulphonates, and in some cases small aromatic or heterocyclic compounds such as imidazolines or dialkylbenzenes. For more complex molecules, in particular natural products, or for mixtures of substances, the definition of appropriate SMILES code fragments did not seem feasible. In several cases, when too many hits were achieved, the structure-related search was combined with use-related parameters such as the sector of use (textile) or the product category (biocides). The overall evaluation of applied search strategies is summarized in Table 10.

Table 10. Frequency of applied search strategies.

	Successful	Not successful	Not performed
Trivial name	68	9	13
Systematic name	47	-	43
SMILES code	20	1	69
Use	11	-	-

Only in one instant, it was not possible to identify any representative substance for a chemical class. This was the case for galactomannan derivatives, where the common name search was unsuccessful and due to the complexity of the structure, no systematic name or SMILES code

fragment could be derived. There were some other cases with only very few search results, which were either also complex natural products, certain synthetic polymers, or very narrowly defined substances classes, for which actually only a small number of representatives exist, such as tributylphosphates. As described above, the search could already be terminated after the initial step, when a sufficient number of results had been obtained, or the results had to be narrowed down because they were too many. In some cases, even larger number of ≥ 100 results were retrieved and reported, when it was possible to define structurally distinct sub-classes. This was particularly the case for some fatty alkyl derivatives, alkyl phosphates or sulphates/sulphonates, halogenated benzenes, and aromatic carboxylic acids and their esters. Table 11 gives an impression of the number of results most typically retrieved.

Table 11. Frequency distribution of the number of valid results retrieved.

No. of valid results	No. of substance classes
0-10	23
11-49	48
50-99	11
≥ 100	8

The database of chemicals maintained by the European Chemicals Agency proved to be a valuable tool for the identification of chemicals based on structure-related search strategies. Many substances are registered under more than one name, and a single registration often contains several synonyms and alternative names. These redundancies make it easier to find substances by more than one approach. However, there also seem to be some short-comings and inconsistencies. In several cases, different spellings such as sul**ph**ate *vs* sul**f**ate gave different results, but there seemed to be no regularity in this, i.e. while many substances could be identified by either spelling variant, some other cases were only retrieved with one particular spelling. Another sensitivity concerns interpunctuation: For instance, different results were obtained for “amine**,** ethoxylated” as compared to “amine ethoxylated”. It is not clear which system of canonical SMILES codes is applied to database entries. While the ECHA guidance for naming and identification of substances gives certain recommendations for publicly available SMILES code generators (ECHA 2017), the SMILES codes that can be found in the database do not actually corresponds to those derived from those recommended generators. Finally, as was already observed during the design of this study, the information regarding the information regarding the use of substances is very incomplete.

The developed structure-based approach for the identification of chemicals that may be potentially relevant for certain technical processes can be applied to lists of individual chemicals if there is a well-established relationship between the chemical structure and the technical function. This could be demonstrated with the help of a case study for an alkyd production site. In other case studies, namely from the surface treatment of metals, the technical functions of the chemicals listed were not clearly described (different and in part conflicting terminologies were used in different case studies within this industry sector). If the definition of technical functions could be unified, better progress might be made with the definition of relevant chemical classes. Application of the structure-based approach for the identification of relevant chemicals is not possible for industries or processes that use inorganic chemicals.

6. References

Chemical Abstracts Service (2007): Naming and Indexing of Chemical Substances for Chemical Abstracts™. https://www.cas.org/sites/default/files/documents/indexguideapp.pdf

Daylight Chemical Information Systems (2011): Daylight Theory Manual. https://www.daylight.com/dayhtml/doc/theory/index.pdf

ECHA (2017): Guidance for identification and naming of substances under REACH and CLP. https://echa.europa.eu/documents/10162/23036412/substance_id_en.pdf/ee696bad-49f6-4fec-b8b7-2c3706113c7d

ECHA (2019): Plastic additives initiative. Supplementary Information on Scope and Methods. https://echa.europa.eu/documents/10162/13630/plastic_additives_supplementary_en.pdf/79bea2d6-8e45-f38c-a318-7d7e812890a1

European Commission (2003): Integrated Pollution Prevention and Control (IPPC). Reference Document on Best Available Techniques for the Textile Industry. https://eippcb.jrc.ec.europa.eu/reference/BREF/txt_bref_0703.pdf

European Commission (2006): Integrated Pollution Prevention and Control (IPPC). Reference Document on Best Available Techniques for the Surface Treatment of Metals and Plastics. https://eippcb.jrc.ec.europa.eu/sites/default/files/2019-11/stm_bref_0806.pdf

European Commission (2007): Reference Document on Best Available Techniques in the Production of Polymers. https://eippcb.jrc.ec.europa.eu/sites/default/files/2019-11/pol_bref_0807.pdf

Fahl, J. (2000): Esteröle und ihre Eigenschaften. DIE KÄLTE & Klimatechnik 4, 40-45.

HAZBREF (2019a): Case Study: Borealis Polymers, Porvoo site, Finland.

HAZBREF (2019b): Case Study: YARA Siilinjärvi and Uusikaupunki sites, Finland.

HAZBREF (2020a): Case study on polymers installation: Production of alkyd resins in Akzo Nobel Baltics AS, Estonia.

HAZBREF (2020b): Case Study: Aurajoki oy, Finland.

HAZBREF (2020c): Case study on STM installation: Surface treatment of metals, Estonia Case 1.

HAZBREF (2020d): Case study on STM installation: Surface treatment of metals, Estonia Case 2.

HAZBREF (2020e): Case Study Report: Surface Treatment of Metals and Plastics, Poland.

International Union of Pure and Applied Chemistry (2019): Compendium of Chemical Terminology. https://goldbook.iupac.org/

Annex: the complete list of substance classes for the textile industry

Green	Search completed	Blue	Individual substance (no search needed)
Grey	Duplicate (see cross-reference)	Red	Search not feasible

Refe-rence No.	Cross-refe-rence	Technical function
1		**Surfactants**
1A		*Non-ionic*
1A.1		Fatty alcohol ethoxylates
1A.2		Fatty acid ethoxylates
1A.3		Alkylphenol ethoxylates (APEOs)
1A.4		Fatty amine ethoxylates
1A.5		Triglyceride ethoxylates
1A.6		Ethylene oxide/propylene oxide adducts
1B		*Anionic*
1B.1		Alkyl sulphonates
1B.2		Alkyl aryl sulphonates
1B.3		Naphthalenesulphonates
1B.4		Ligninsulphonates
1B.5		Alkyl sulphates
1B.6		Alcohol ethoxysulphates
1B.7		Sulphated alkanolamides
1B.8		Dialkylsulphosuccinates
1B.9		Sulphated/sulphonated vegetable oils
1B.10		Alkyl carboxylates (e.g. sodium palmitate, -stearate)
1B.11		Alkylether phosphates
1C		*Cationic*
1C.1	6C.10, 6C.11	Quaternary ammonium compounds (salts)
1D		*Amphoteric*
1D.1		Betaine derivatives
1D.2		Imidazolines
1D.3	1A.4	Modified fatty amino ethylates
2		**Auxiliaries and finishing agents for fibre and yarn manufacturing**
2A		*Lubricants*
2A.1		Mineral oils
2A.2		Ester oils
		Synthetic lubricants:
2A.3		Polyalphaolefins (PAO)
2A.4		Dialkyl benzenes

Refe-rence No.	Cross-refe-rence	Technical function
2A.5	2A.2	Synthetic esters
2A.6		Polyglycols
2A.7		Silicones
2A.8	1A.6	EO/PO adducts
2B		*Emulsifiers*
2B.1	1A.1	Ethoxylated fatty alcohols
2B.2	1A.2	Ethoxylated fatty acids
2B.3		Ethoxylated sorbitan esters
2B.4	1A.3	Alkyl phenol ethoxylates (APEO)
2B.5	1A.5	Partial glycerides and ethoxylated triglycerides
2B.6		Fatty amides
2B.7	1B.9	Sulphonated and sulphated vegetable oils
2C		*Wetting agents*
2C.1	2D.2	Short-chain alkyl phosphates
2D		*Antistatic agents*
2D.1	1B	Anionic surfactants
2D.2		Alkyl phosphates
2D.3		Sarcosides
2D.4		Amine oxides
2D.5	1B.8	Sulpho succinates
2E		*Biocidal additives*
2E.1		Formaldehyde-containing compounds
2E.2		Imidazolinones
2E.3		Isothiazolinones
3		**Sizing agents**
3A		Starch
3B		*Starch derivatives:*
3B.1		Dextrins
3B.2		Starch esters (most common: phosphates, acetates)
3B.3		Starch ethers (most common: hydroxyethyl, hydroxypropyl, carboxymethyl)
3C		Cellulose derivatives (carboxymethly cellulose)
3D		Galactomannan derivatives (hydroxypropyl and carboxymethyl galactomannan)
3E		*Polyvinyl alcohol (PVA)*
3E.1		Partially hydrolysed PVA (ca. 88%)
3E.2		Fully hydrolysed PVA (ca. 98%)
3E.3		Co-polymers of PVA with methyl methacrylate or other carboxyl monomers
3F		*Poly(meth)acrylates*
3F.1		Considerable variation in chemical structure (building blocks: acrylic acid, acrylic esters, acrylamide, acrylonitrile, methacrylic acid)

Refe-rence No.	Cross-refe-rence	Technical function
3G		Polyesters
3G.1		Condensates of aromatic dicarboxylic acids with diols (e.g. ethylene gylcol, diethylene glycol) and sulphonated aromatic dicarboxylic acids
4		**Detergents/wetting agents**
4A	1A	Non-ionic surfactants
4B	1B	Anionic surfactants
5		**Auxiliaries containing sequestering agents**
5.1		Ethylenediamine tetraacetate (EDTA)
5.2		Nitrilotriacetate (NTA)
5.3		Diethylenetriamine pentaacetate (DTPA)
5.4		Phosphonic acid derivatives (phosphonates)
5.5		Gluconic acid derivatives (gluconates)
5.6	3F	Polyacrylates
6		**Dyeing auxiliaries**
6A		*Wetting, penetrating and de-aerating agents*
	1	*Surfactants, e.g.:*
6A.1	1A.1	Alcohol polyglycol ethers
6A.2	1A.2	Polyglycol esters
6A.3	1B.1	Alkane sulphones
6A.4	1A.4	Ethoxylated amines
6B		*Dispersing agents*
6B.1		Condensation products of naphthalene sulphonic acid with formaldehyde
6B.2	1B.4	Lignosulphonates
6B.3	1B	Anionic surfactants
6B.4	1A	Non-ionic surfactants
6C		*Levelling agents*
		Surfactants, e.g.:
6C.1	1A.1	Fatty alcohol ethoxylates
6C.2	1A.4	Fatty amine ethoxylates
6C.3	1A.2	Fatty acid ethoxylates
6C.4	1A.3	Alkyl phenol ethoxylates (APEO)
6C.5	1A.6	EO/PO adducts
6C.6	1B.5	Fatty alcohol sulphates
6C.7		Alkyl aryl sulphates
6C.8		Polyamide amines
6C.9		Polyvinylpyrrolidone
		Quaternary ammonium compounds, e.g.:
6C.10		Quaternary ammonium salts with C12-C14 fatty alkyl side chains
6C.11		Quaternary ammonium salts with aromatic ring systems

Reference No.	Cross-reference	Technical function
6C.12		Ethoxylated castor oil
6C.13		Ethoxylated stearic acid
6C.14		Mixtures of alcohols, esters or ketones of medium chain length with emulsifying systems
6D		*Acid donors*
6D.1	7D.2	Organic acid esters
6D.2	1A.1	Fatty alcohol ethoxylates
6D.3	1B.2	Alkyl aryl sulphonates
6E		*Antifoaming agents*
6E.1	2A.7	Based on silicone derivatives
6E.2		Tributylphosphates
6E.3	2D.2	Phosphoric esters
6E.4		High molecular-weight alcohols
6E.5		Fluorine derivatives
6F		*Carriers*
6F.1		Halogenated benzenes (1,2-dichlorobenzene, 1,2,4-trichlorobenzene, dichlorotoluene)
6F.2		Aromatic hydrocarbons such as alpha- and beta-methylnaphthalene, diphenyl, trimethyl benzene etc.
6F.3		Phenols such as o-phenylphenol, benzylphenol etc.
6F.4		Aromatic carboxylic acids and their esters such as methyl, butyl and benzyl benzoate, methylsalicylate, phthalic acid, dimethylphthalate, dibutylphthalate, diethylhexylphthalate
6F.5		Alkyl phthalimides such as N-butylphthalimide
7		**Printing auxiliaries**
7A		*Thickening agents*
7A.1	3A-3D	Natural polysaccharides, unmodified or modified (e.g. seed derivatives, starch degradation products, alginates)
7A.2	3F	Synthetic polymers based on polyacrylic acid
7A.3	2A.1	Mineral oils
7A.4		White spirit (mixture of aliphatic hydrocarbons with C12-C50 chain length)
7B		*Binders*
7B.1		Polymer dispersions based on acrylates, butadiene, or vinyl acetate
7C		*Fixing agents*
7C.1		Melamine-formaldehyde condensates
7D		*Plasticisers*
7D.1	2A.7	Silicones
7D.2		Fatty acid esters
7E		*Emulsifiers*
		Non-ionic surfactants, e.g.:

Refe-rence No.	Cross-refe-rence	Technical function
7E.1	1A.3	Aryl polyglycol ethers
7E.2	1A.1	Alkyl polyglycol ethers
8		**Finishing auxiliaries**
8A		*Cross-linking agents*
8A.1		Agent based on formaldehyde and hydroxymethyl urea or methoxymethyl urea
8A.2	7C.1	Agents based on hydroxymethyl melamine or methoxymethyl melamine
8A.3		Agents based on derivatives of hydroxymethyl dihydroxyethylene urea
8A.4		Agents based on dimethyl urea and glyoxal
8B		*Catalysts*
8B.1		Magnesium chloride
8B.2		Organic acids (e.g. alpha-hydroxycarboxylic acids)
8B.3		Inorganic Lewis acids
8B.4		Ammonium salts (e.g. chloride, sulphate and nitrate)
8C		*Additives (softeners, stiffening agents, etc.)*
8C.1	3F	Polyacrylates
8C.2	2A.7	Polysiloxanes
8C.3		Polyethylene waxes (partially oxidised polyethylenes)
8C.4		Polyurethanes
8C.5		Fatty acid condensation products (cationic softeners)
8C.6		Polyvinyl acetate
8D		*Biocides*
		Moth-proofing agents:
8D.1		Permethrin
8D.2		Cyfluthrin
8D.3		Sulcofuron
		Antimicrobials or odour suppressants:
8D.4		Zinc organic compounds
8D.5		Tin organic compounds
8D.6		Dichlorophenyl(ester) compounds
8D.7		Benzimidazole derivatives
8D.8		Triclosan
8D.9	2E.3	Isothiazolinones
8E		*Antistatic agents*
8E.1	1C.1	Quarternary ammonium compounds
8E.2	2D.2	Alkyl phosphates
8E.3	1B.11	Alkylether phosphates
8F		*Flame-retardants*
		Inorganic compounds:
8F.1		Diammonium phosphate

Refe-rence No.	Cross-refe-rence	Technical function
8F.2		Aluminum sulphate
8F.3		Ammonium sulphate
8F.4		Aluminum hydroxide
8F.5		Titanium halogenated salts
8F.6		Zirconium halogenated salts (e.g. potassium hexafluorozirconate)
		Halogenated compounds:
8F.7		Polybrominated diphenyl ethers (PBDE), e.g. pentabromodiphenyl ether (penta-BDE), octabromodiphenyl ether (octa-BDE), decabromodiphenyl ether (deca-BDE)
8F.8		Polybromo biphenyls (PBB), e.g. decabromobiphenyl
8F.9		Tetrabromobisphenol A (TBBA)
8F.10		Hexabromocyclodecane
8F.11		Chloroparaffins
		Phosphor-organic compounds:
8F.12	2D.2	Phosphate esters
8F.13		Phosphine oxides
8F.14		Phosphoric acids
8F.15		Phosphonic acid (2-((hydroxymethyl)carbamyl)ethyl)-dimethyl ester
8F.16		Phosphonium salt and urea precondensates
8G		*Repellents*
8G.1		Wax-based repellents, typically containing paraffin (25%) and zirconium or aluminium salts (5-10%)
8G.2		Resin-based repellents (condensation products of fatty acids, alcohols or amines with methylolated melamines)
8G.3		Silicone repellents, consisting of (modified) dimethylpolysiloxane, emulsifiers, glycols and water
8G.4		Fluorochemical repellents, mostly copolymers of fluoroalkyl acrylates and methacry-lates
8H		*Softeners*
8H.1	1A	Non-ionic surfactants
	1C	*Cationic surfactants:*
8H.2	1C.1	Quaternary ammonium compounds, such as stearyl or distearyl dimethyl ammonium chloride
8H.3		Amido amines formed by reaction of a fatty acid or a glyceride and a short chain poly-amine
8H.4		Imidazolines which can be acetylated or reacted with ethylene oxide
8H.5	8C.3	Polyethylene waxes
8H.6	2A.7	Silicones
9		**Coating compounds and auxiliaries**
9A		*Coating powders and pastes*
9A.1	8C.3	Polyolefins (e.g. polyethylene)
9A.2		Polyamide 6
9A.3		Polyamide 6.6

Refe-rence No.	Cross-refe-rence	Technical function
9A.4		Copolyamides
9A.5	3G	Polyester
9A.6	8C.4	Polyurethane
9A.7		Polyvinylchloride
9A.8		Polytetrafluoroethylene
9A.9		Ethylene vinyl acetate
9A.10		Bitumen
9A.11		x-SBR latex
9B		*Polymer dispersions (aqueous formulations)*
9B.1	3F	Poly(meth)acrylate (butyl, ethyl, methyl etc.)
9B.2	3F	Polyacrylic acid
9B.3	3F	Polyacrylonitrile
9B.4	3F	Polyacryloamide
9B.5		1,3-polybutadiene
9B.6		Polystyrene
9B.7	8C.4	Polyurethane
9B.8	9A.7	Polyvinylchloride
9B.9	8C.6	Polyvinylacetate
9B.10		Copolymeres of the above-mentioned polymers
9C	7C.1	Melamine resins produced by reaction of melamine and formaldehyde
9D		*Polymer dispersions (organic solvent-based formulations)*
9D.1	8C.4	Polyurethanes
9D.2	2A.7	Silicones
10		**Solvents**
		Trichloroethylene (TCE)
		Benzene
		White spirit
		Solvent naphtha
		Perchloroethylene
		Glycol ethers (e.g. dipropylene glycol tertiary-butyl ether)
		Liquid silicone (decamethylcyclopentasiloxane or D5)
		Liquid CO2
11		**Other chemicals**
		Sodium hydroxide
		Sodium carbonate
		Sulphuric acid
		Hydrochloric acid
		Aluminium chloride
		Sodium hypochlorite
		Sodium dichloroisocyanuarate (DCCA)

Refe-rence No.	Cross-refe-rence	Technical function
		Chlorine
		Peroxysulphate
		Permanganate
		Bisulphate