

RISK *Identifizieren | Bewerten* *Handeln | Kommunizieren* IDENT

BMBF-Verbundprojekt RISK-IDENT:

STOFF-IDENT: Datenumfang und –qualität
Dr. Marion Letzel, LfU

gefördert vom:



Bundesministerium
für Bildung
und Forschung



Bayerisches Landesamt für
Umwelt



HOCHSCHULE
WEIHENSTEPHAN-TRIEDSDORF
UNIVERSITY OF APPLIED SCIENCES



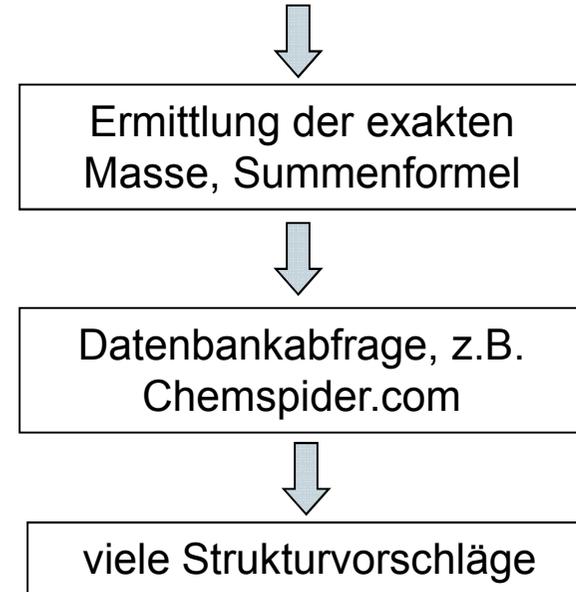
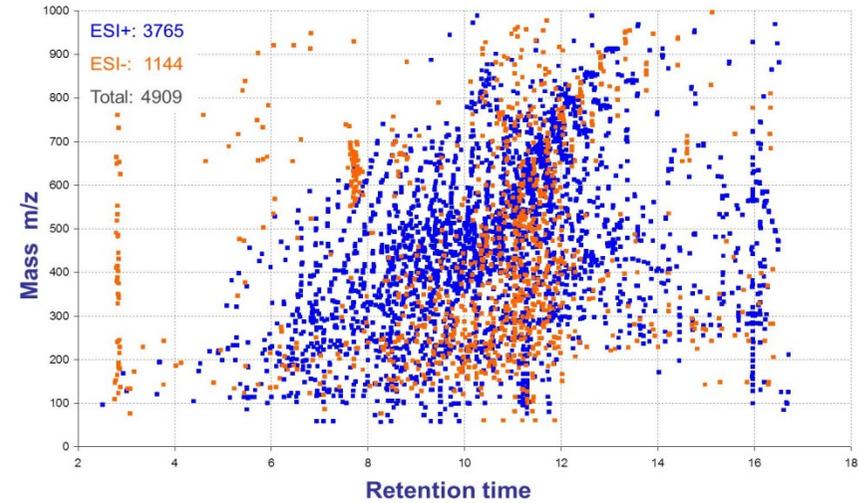
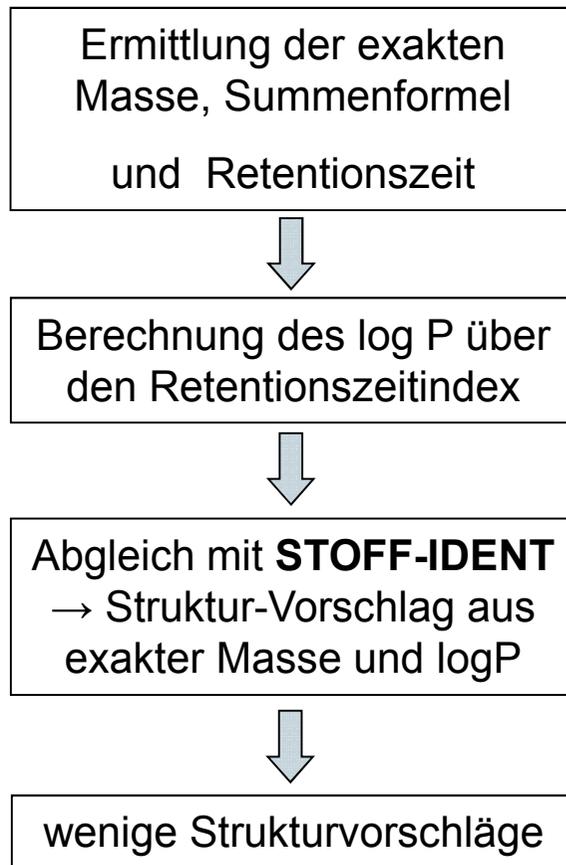
TUM
Technische Universität München

Zweckverband
Landeswasserversorgung



CONDIAS
CONDUCTIVE DIAMOND PRODUCTS

Vorgehensweise Identifizierung



STOFF-IDENT: Datenumfang und -qualität

STOFF-IDENT: Datenbank gewässerrelevanter Stoffe

- Industriechemikalien (v.a. REACH-registrierte Stoffe)
- Humane Arzneimittelwirkstoffe und bekannte Metabolite
- (Tierarzneimittel)
- Pflanzenschutzmittel und -Metabolite
- Biozide
- Weitere (bisher nachgewiesene) Stoffe
- Transformationsprodukte

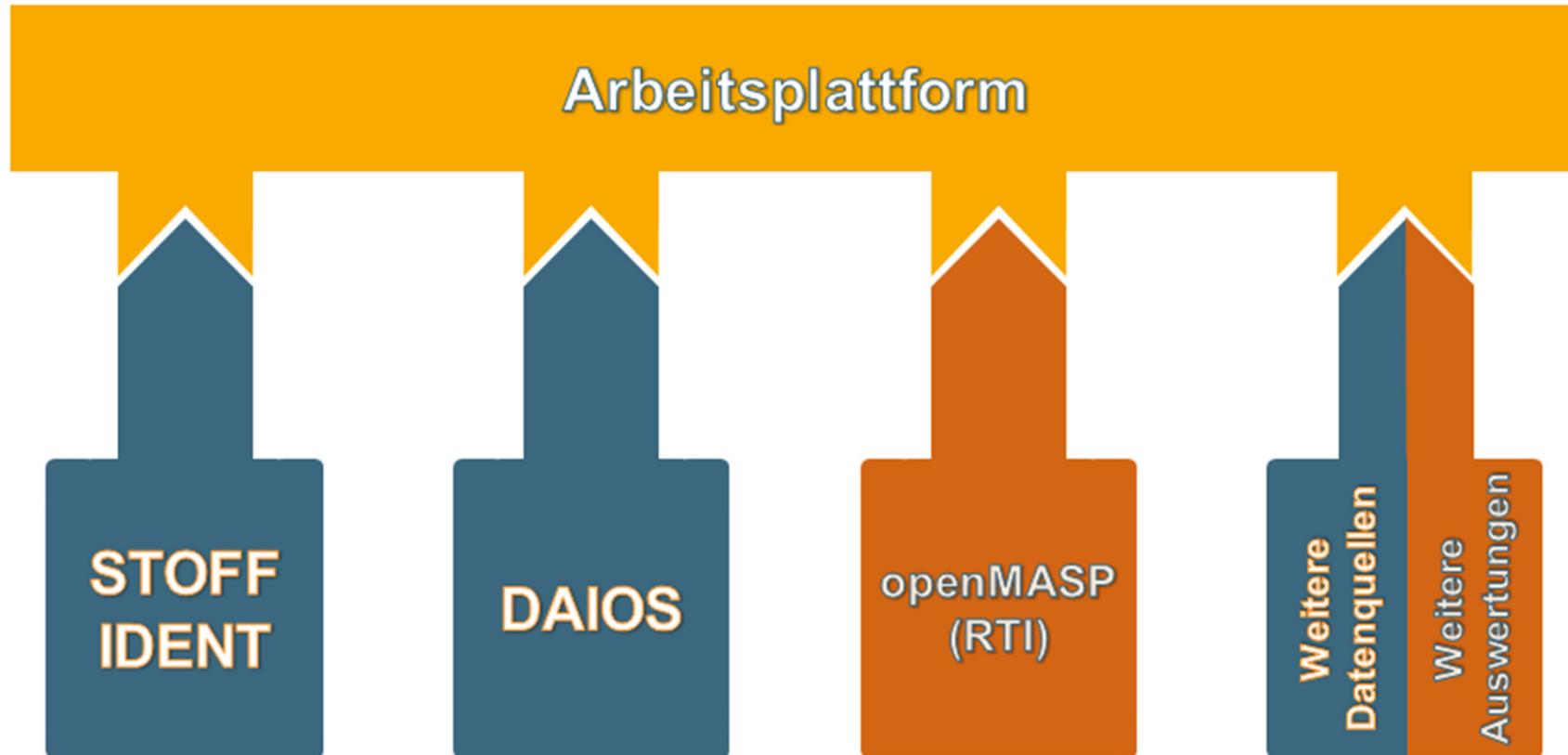


Fotos: LfU

STOFF-IDENT: Datenattribute

- Name
- CAS-Nr.
- EC-Nr.
- SMILES
- Chem. Name (IUPAC)
- Summenformel
- Exakte, monoisotopische Masse
- Log P_{OW}
- log D_{OW} bei pH 3, 5, 7, 9
- Mengenband
- Zusatzinformationen wie pKa, Siedepunkt, Wasserlöslichkeit
- (Stoffgruppe, Anwendung)

Einbindung der Datenbank STOFF-IDENT



REACH

Registrierungsfristen:

- 1.12.2010: Stoffe >1000 t/a
+ R50/53 >100 t/a
+ CMR (Category 1+2) ≥ 1 t/a
- 1.06.2013: Stoffe >100 t/a
- 1.06.2018: Stoffe 1-100 t/a

- 7.12.2010: 2.992 Registrierungen
- 1.06.2013: 6.600 Registrierungen
- 3.12.2013: 11.766 Registrierungen
- Manuelle Entfernung von anorganischen Stoffe, Reaktions- und Stoffgemischen sowie Komplexen
- Bestand in STOFF-IDENT: ca. 4.350 Stoffe



REACH-Stoffe: Beispiel

EC / List No.	CAS No.	Name	Total Tonnage Band
700-853-5	943586-12-7	16-[[[(1S)-1-carboxy-4-(2,5-dioxopyrrolidin-1-yl)oxy-4-oxo-butyl]amino]-16-oxo-hexadecanoic acid	1 - 10 tonnes per annum
700-854-0	256473-04-8	2-Chloro-3-(2,2,2-trifluoroethoxy)pyridine	Intermediate Use Only
700-855-6		Sulfonation products of disodium sulfite with (esterification products of C9-11 (branched and linear) Alkyl Polyglycosides with maleic anhydride)	100 - 1,000 tonnes per annum
700-856-1	1247119-31-8	(R,R)-1,6-diphenylhexane-2,5-diamine dihydrochloride	Intermediate Use Only
700-856-1	1247119-31-8	2,5-Hexanediamine, 1,6-diphenyl-, hydrochloride, (2R,5R)-	Intermediate Use Only
700-857-7		Reaction mass of 2-hydroxy-4-(methylthio)butanoic acid and 2-hydroxy-4-(methylthio)butanoic acid dimer	Intermediate Use Only
700-869-2		Solid Recovered Fuel (Municipal Solid Waste and similar wastes, processed)	10,000 - 100,000 tonnes per annum
700-872-9		Not yet assigned	Intermediate Use

Humane Arzneimittelwirkstoffe

- Quelle IMS MIDAS ®
2009/Umweltbundesamt
- 1517 small molecule-Wirkstoffe, aussortiert wurden u.a. Pflanzen, Tiere, Mikroorganismen, Metalle, Vitamine, Nukleine, Mineralien, Proteine/Peptide/AS, Diagnostika außer Röntgenkontrastmittel
- Stoffe mit CAS-Nummern: 440
- Manuelles Einfügen der restlichen CAS-Nummern
- Bestand in STOFF-IDENT: 1.517 Arzneimittelwirkstoffe



Pflanzenschutzmittel und ihre Metabolite

- Liste der 2014 in D zugelassenen PSM-Wirkstoffe (239 Stoffe, Quelle: BVL)
- Liste PSM-Metabolite über $1\mu\text{g/L}$ in Lysimeterversuchen (50 Stoffe, Quelle: BVL)
- Reemtsma et al. 2013: Metabolite aus Grund- und Oberflächenwasser: 105 Stoffe, teilweise mit CAS, Summenformel
- Manuelles Einfügen der restlichen CAS-Nummern



Biozide

- zugelassene Biozide (EU),
Altwirkstoffe: Commission Regulation
(EC) Nr. 1451/2007 (Anhang I): 583
Stoffe mit CAS
- zu prüfende Wirkstoffe (Anhang II):
234 Stoffe mit CAS, Überschneidung
mit Anhang I
- Aufnahme in Anhang I der Biozid-
Richtlinie 98/8/EG: 63 Stoffe mit CAS
- Alle Listen bearbeitet: Gemische,
Anorganik, Mikroorganismen,
natürliche Öle & Extrakte entfernt



Weitere Stoffe / Stofflisten (Auswahl)

Stoffsammlung aus Analytik-Laboren und NORMAN

- LW: 2665 Stoffe aus GC- & LC-MS-Messungen
- Forensik Innsbruck: 1207 Stoffe aus LC-MS-Messungen
- NORMAN-list of emerging substances (694); NORMAN-list of candidate emerging substances (180)
- LfU: 94 Stoffe

Stofflisten Literatur

- Howard & Muir 2010 + 2013: Verunreinigungen, Byprodukte, 930 Stoffe
- Field et al. 1994: secondary alkane sulfonates
- Drewes et al. 2008: abwasserbürtige Stoffe, 62 Stoffe
- HAM: TPs von Li et al. 2014, Kern et al. 2010, RISK-IDENT
- Drugs: Frost & Griffiths 2008, EMFDDA Selected Issue 2010

► *Quellen für Eintragungen in STOFF-IDENT sind in der Datenbank hinterlegt*

Automatisches Einlesen mit dem SI-Crawler

Target: SuMS-Liste_LW
2665 records - 2665 visible

Completen...	Suspicious	Name	Source	CAS	Source	SMILES	Source	IUPAC	Source	Formula
●	●	Mirex	Zweckver	2385-85-5	Zweckverband L					
●	●	Bromfenvinphos-methyl	Zweckver	13104-21-7	Zweckverband L					
●	●	Prohexadione-calcium	Zweckver	127277-53-6	Zweckverband L					
●	●	2,4-DB	Zweckver	94-82-6	Zweckverband L					
●	●	Chlorthenoxazine	Zweckver	132-89-8	Zweckverband L					
●	●	Sulfamethoxazol, Natriumsalz	Zweckver	4563-84-2	Zweckverband L					
●	●	Metamitron	Zweckver	41394-05-2	Zweckverband L					
●	●	Sulfadiazine	Zweckver	68-35-9	Zweckverband L					
●	●	Tetralinone	Zweckver	529-34-0	Zweckverband L					
●	●	2-Acetylacetophenone	Zweckver	704-00-7	Zweckverband L					
●	●	Dimethylphthalate (DMP)	Zweckver	131-11-3	Zweckverband L					
●	●	Brallobarbital	Zweckver	561-86-4	Zweckverband L					
●	●	Chlortoluron benzoic acid (CGA 1	Zweckver	59587-01-8	Zweckverband L					
●	●	Acetamidiprid	Zweckver	135410-20-7	Zweckverband L					
●	●	Clofibric acid (metabolite of CLOI	Zweckver	882-09-7	Zweckverband L					
●	●	Mecoprop (MCP)	Zweckver	7085-19-0	Zweckverband L					
●	●	Mecoprop-p (MCP-P)	Zweckver	16484-77-8	Zweckverband L					

- Anzeigen der Vollständigkeit mit Farbampel
- Vervollständigen der Daten mit chemicalize.org
- Berechnung der monoisotopischen Masse aus der Summenformel

Prüfregeln beim automatischen Einlesen

- Substanz muss CAS-Nummer oder SMILES-Code haben
- CAS- Nummer: Prüfziffer wird überprüft
- $-20 < \log P < 20$
- Überprüfung: passt die Summenformel zum SMILES-Code?
- Überprüfung: CAS, SMILES und IUPAC bei doppelten Datensätzen

Completen...	Suspicious	Name	Source	CAS	Source	SMILES
●	●	Aciclovir	UBA	59277-89-3	UBA	<chem>Nc1nc(=O)c2ncn(COCCO)c2[nH]1</chem>
●	●	ACIPIMOX	UBA	51037-30-0	Wikipedia	<chem>Cc1cncc(C(O)=O)[n+]1[O-]</chem>
●	●	ACITRETIN	UBA	55079-83-9	Wikipedia	<chem>COc1cc(C)c(\C=C\C(\C)=C\C=C\C(\C</chem>
●	●	ACLARUBICIN	UBA	57576-44-0	Wikipedia, chem	<chem>CC[C@@]1(O)C[C@H](OC2CC(C(OC3</chem>
●	●	Acriflavinium Chloride	UBA	8063-24-9	chemicalbook.cc	<chem>[Cl-].Nc1ccc2cc3ccc(N)cc3nc2c1.Cc1c</chem>
●	●	ACTINOQUINOL	UBA	15301-40-3	http://www.drug	<chem>CCOc1ccc(c2cccnc12)S(O)(=O)=O</chem>
●	●	ADAPALENE	UBA	106685-40-9	Wikipedia; chem	<chem>COc1ccc(cc1C12CC3CC(CC(C3)C1)C2)</chem>
●	●	Adefovirdipivoxil	UBA	142340-99-6	UBA	<chem>CC(C)(C)C(=O)OCOP(=O)(COCCn1cn</chem>
●	●	Ademetionine disulfate tosylate	UBA	97540-22-2	chemicalbook.cc	<chem>OS([O-])(=O)=O.[O-]S([O-])(=O)=O.Cc</chem>
●	●	ADRENALONE	UBA	99-45-6	chemicalbook.cc	<chem>CNCC(=O)c1ccc(O)c(O)c1</chem>
●	●	Agomelatin	UBA	138112-76-2	UBA	<chem>COc1ccc2cccc(CNCC(C)=O)c2c1</chem>
●	●	AJMALICINE	UBA	483-04-5	chemicalbook.cc	<chem>COC(=O)C1=CO[C@@H](C)[C@H]2Cl</chem>
●	●	AJMALINE	UBA	4360-12-07	chemicalbook.cc	
●	●	ALATROFLOXACIN	UBA	157605-25-9	wikipedia, chemi	<chem>CS(O)(=O)=O.C[C@H](N)C(=O)N[C@H]</chem>
●	●	ALBENDAZOLE	UBA	54965-21-8	Wikipedia	<chem>CCCSc1ccc2nc(NC(=O)OC)[nH]c2c1</chem>
●	●	ALCLOMETASONE	UBA	66734-13-2	Wikipedia	
●	●	ALCLOXA	UBA	1317-25-5	chemicalbook.cc	

Fehlerbehebung (in Arbeit)

- Korrektur der gelb markierten inkonsistenten Datensätze („suspicious“)
- Salze, Addukte entfernen bzw. in Säuren umwandeln (manuell)
- Automatischer Fehlerabgleich bei doppelten Datensätzen

Output: Target_T1-bearbeitet_SMILES_IUPAC ☒

109 records - 109 visible

type to filter Quick filter -- Show all --

Completen...	Suspicious	Name	Source	CAS	Source	EC number	Source	SMILES	Source	IUPAC	Source
●	●	2,6-xylenol	REACH	576-26-1	REACH	209-400-1	HAM	Cc1cccc(C)c1O	www.chemic	2,6-dimethylpl	www.c
●	●	triphenylphosphine oxide	REACH	791-28-6	REACH	212-338-8	HAM	O=P(c1ccccc1)(c1ccccc1)c1ccccc1	www.chemic	(diphenylphos	www.c
●	●	polychloro copper phthalocyanin	REACH	1328-53-6	REACH	215-524-7	HAM	Clc1cc(Cl)c(Cl)c(Cl)c1	www.chemic	3,4,5,6,12,14,15	www.c
●	●	tris(methylphenyl) phosphate	REACH	1330-78-5	REACH	215-548-8	HAM	Cc1ccc(OP(=O)(c1ccccc1)c1ccccc1)cc1	www.chemic	tris(4-methylpl	www.c
●	●	2-Butanone, peroxide	REACH	1338-23-4	REACH	700-954-4	REACH	CCC(C)(OO)OO	www.chemic	2-[(2-hydroper	www.c
●	●	2-ethylhexyl 4-methoxycinnamat	REACH	5466-77-3	REACH	226-775-7	HAM	CCCCC(CC)COC	www.chemic	2-ethylhexyl (2	www.c
●	●	octocrilene	REACH	6197-30-4	REACH	228-250-8	HAM	CCCCC(CC)COC	www.chemic	2-ethylhexyl 2-	www.c

Output: Gesamt10_logD ☒

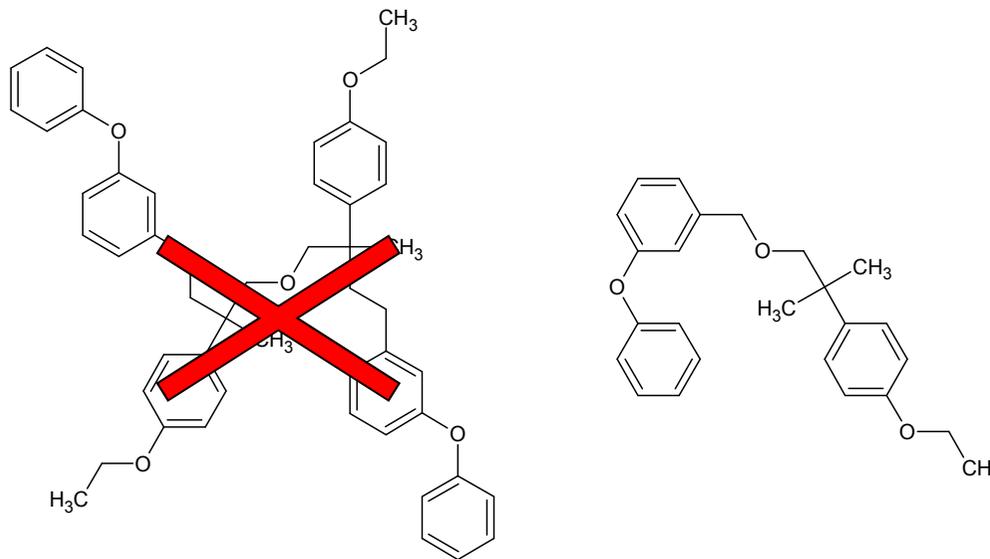
8817 records - 1 visible

Cc1cccc(C)c1O Quick filter -- Show all --

Completen...	Suspicious	Name	Source	CAS	Source	EC number	Source	SMILES	Source	IUPAC	Source
●	●	xylenol	COMMIS	1300-71-6	COMMISSION RE	215-089-3	COMMISSION RE	Cc1cccc(C)c1O	www.chemic	2,6-dimethylpl	www.cher

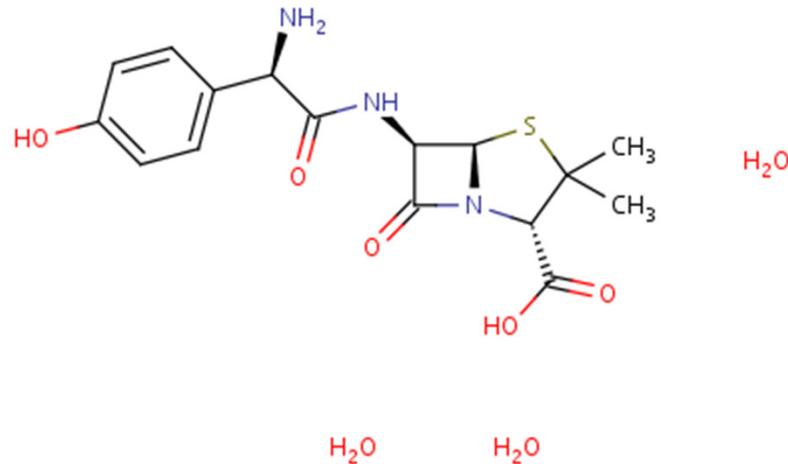
Beispiel: Etofenprox (Insektizid)

- Aus 3 Listen: REACH (nur Name und EC-Nummer), Biozid (Name, EC-Nummer und CAS-Nummer) und PSM (nur Name und CAS-Nummer)
- Vervollständigung der fehlenden Daten mit chemicalize.org
- Abgleich EC-Nummern führt zu Fehlermeldung: SMILES, IUPAC und Masse unterschiedlich
- Fehler bei chemicalize.org



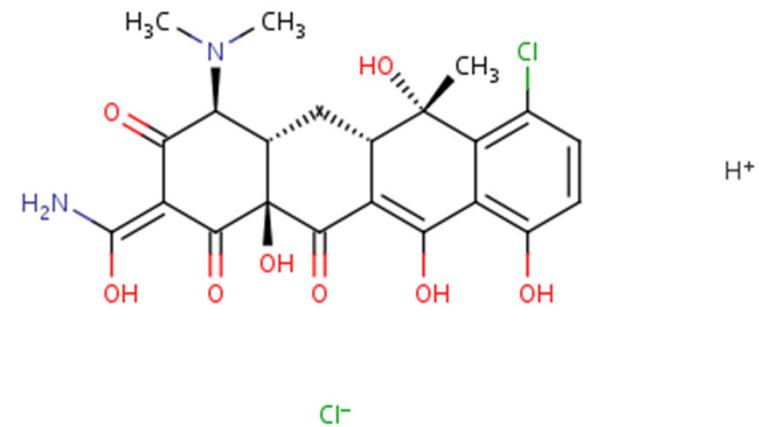
Etofenprox =
3-phenoxybenzyl-2-
(4-ethoxyphenyl)-2-
methylpropyl ether
EC: 407-980-2

Beispiel: Amoxicillin, Chlortetracyclin



CAS-Nummer führt zu: C₁₆H₂₅N₃O₈S
(2S,5R,6R)-6-[(2R)-2-amino-2-(4-hydroxyphenyl)acetamido]-3,3-dimethyl-7-oxo-4-thia-1-azabicyclo[3.2.0]heptane-2-carboxylic acid **trihydrate**; M: 419,1362

Richtig: C₁₆H₁₉N₃O₅S, M: 365,1045



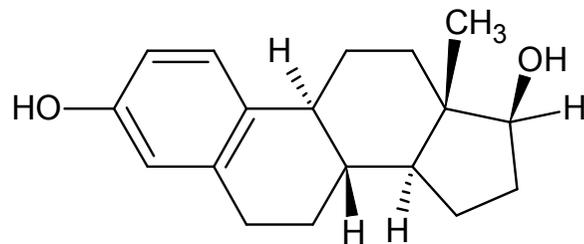
CAS-Nummer führt zu: C₂₂H₂₄Cl₂N₂O₈
(2S,5R,6R)-hydrogen (2Z,4S,4aS,5aS,6S,12aS)-2-[amino(hydroxy)methylidene]-7-chloro-4-(dimethylamino)-6,10,11,12a-tetrahydro-6-methyl-1,2,3,4,4a,5,5a,6,12,12a-decahydrotetracene-1,3,12-trione **chloride**; M: 514,0909

Richtig: C₂₂H₂₃ClN₂O₈, M: 478,1143

Beispiel: 17-beta Estradiol

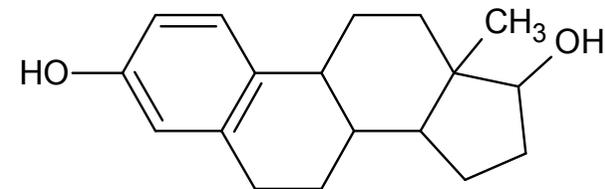
- Stereoselektive SMILES-Codes
- Stereoselektive IUPAC-Namen

C[C@]12CC[C@H]3[C@@H](CCc4cc(O)ccc34)[C@@H]1CC[C@H]2O



(1S,10R,11S,14S,15S)-15-methyltetra
cyclo[8.7.0.0^{2,7}.0^{11,15}]heptadec
a-2(7),3,5-triene-5,14-diol

CC12CCC3C(CCc4cc(O)ccc34)C1CCC2O



13-Methyl-6,7,8,9,11,12,14,15,16,17-deca
hydrocyclopenta[a]phenanthrene-3,17-diol

Zusammenfassung

- Die wesentlichen gewässerrelevanten Stoffgruppen wurden integriert
- Trotz automatischen Einlesetool großer manueller Aufwand der Datenerhebung und Fehlerbeseitigung
- Es befinden sich noch etliche Fehler und unvollständige Datensätze in STOFF-IDENT – work in progress!
- Bitte um Mitteilung von gewässerrelevanten Einzelstoffen oder entsprechenden Literaturstellen (insbesondere neu identifizierte Transformationsprodukte)
- STOFF-IDENT wird ab 2015 am LfU weitergeführt und zumindest um die neu angemeldeten REACH-Chemikalien ergänzt (ca. halbjährliches update)

Projektteam Februar 2014



Vielen Dank für die Aufmerksamkeit!

gefördert vom:



Bundesministerium
für Bildung
und Forschung

