



Rijksinstituut voor Volksgezondheid
en Milieu
*Ministerie van Volksgezondheid,
Welzijn en Sport*

Overview of CADASTER: project goals, achievements and lessons

Willie Peijnenburg



CADASTER

- **General: REACH, ITS**
- **CADASTER: Overview aims/goals**
- **Achievements**
- **Some lessons learned**



REACH

Registration, Evaluation, Authorisation and Restriction of Chemicals

- REACH requires demonstration of safe manufacture and use of chemicals
- REACH based on precautionary principle, aims at achieving proper balance between social, economic and environmental objectives
- **REACH aims to optimise the use of scarce and scattered info on substances**
- **REACH aims to minimise animal testing by optimal use of info on “related” compounds**

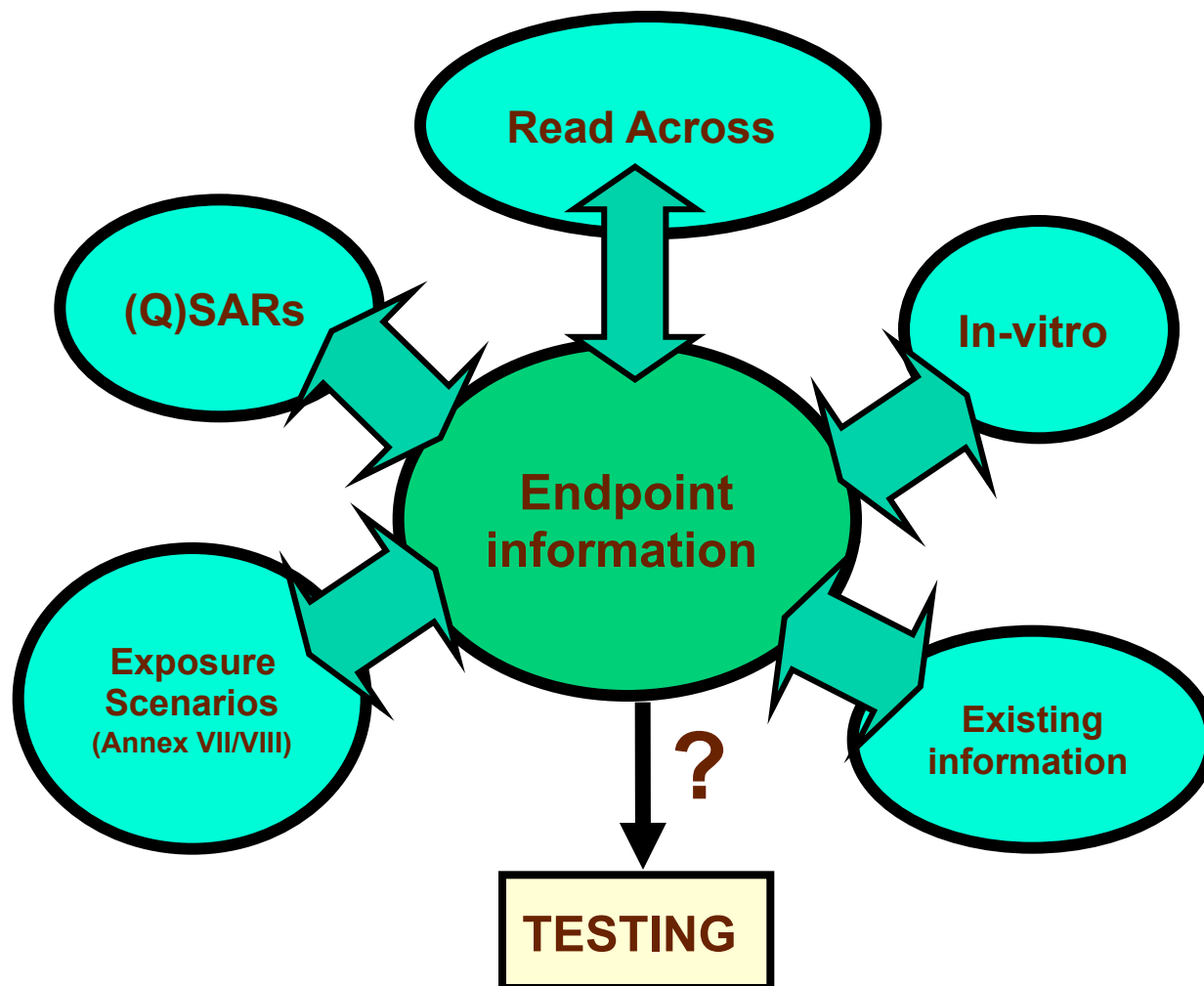


Minimised animal use:

- 1 – Use of validated *in silico* techniques: (Q)SAR/(Q)SPR
- 2 – New *in vitro* test methods, *in vivo* info analogues
- 3 – Minimization actual numbers of animals used, and replacement of animal tests by alternative methods
- 4 – Substance Information Exchange Forums (SIEFs) for obligatory provision of data and cost sharing
- 5 - Requirement of official sanctioning of proposals for tests for compounds with production volumes of above 100 tonnes to minimize animal testing



Intelligent Testing Strategies (ITS)





Goals

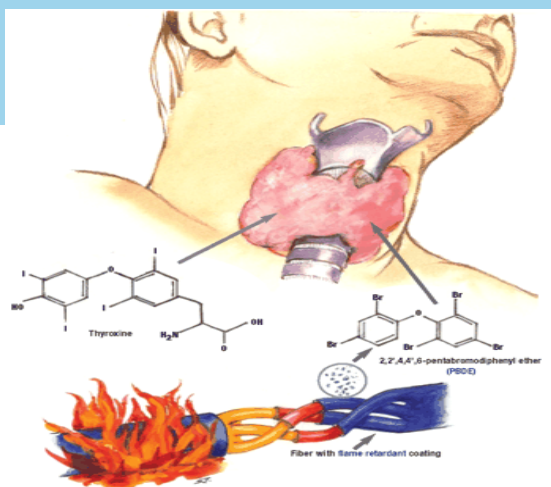
- Exemplify the integration of information, models, strategies for safety, hazard, and risk assessment for 'large' numbers of substances
- Carry out "real" hazard assessment for 'large' numbers of substances according to the basic philosophy of REACH: < costs, animal testing, time
- Thus: Exemplifying how to increase non-testing information whilst quantifying and reducing uncertainty



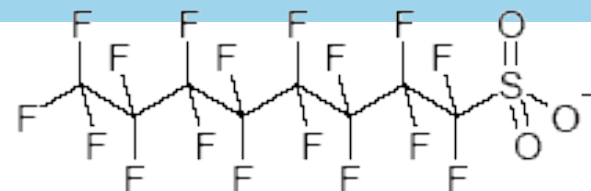
Aim

Provide full environmental hazard and risk assessment according to the REACH philosophy for chemicals belonging to 4 classes of emerging chemicals:

- 1 – Polybrominated diphenylethers (PBDE), typically class of **hydrophobic** chemicals that pose a threat to man and the environment.
- 2 – Perfluoroalkylated substances and their transformation products, like perfluoroalkylated sulfonamides, alkanolic acids, sulfonates. **Persistent hydrophilic compounds** that may be **toxic** for man and environment.
- 3 – Substituted musks/fragrances; a **heterogenic** group of chemicals of varying composition like substituted benzophenones, polycyclic musks, terpene derivatives. **Common emission pattern** in the environment.
- 4 – Triazoles/benzotriazoles: **increasingly used** as pesticides and anti-corrosives.



PBDE: "The PCB's of the future"

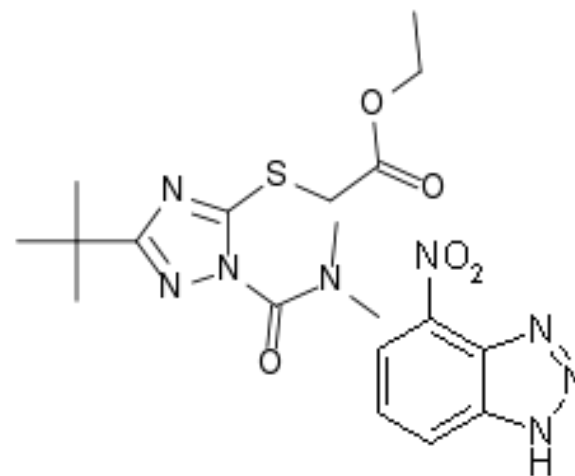


Classification of PFOS-compounds in 22 categories according to OECD

Table 3. Personal care products produced in Germany (1993).

Product category	Tons produced
Bath additives	162,300
Shampoos, hair tonic	103,900
Skin care products	75,500
Hair sprays, setting lotions, hair dyes	71,000
Oral hygiene products	69,300
Soaps	62,600
Sun screens	7,900
Perfumes, aftershaves	6,600
Total	559,100

Pharmaceuticals and Personal Care Products in the Environment: Agents of Subtle Change?



Microbicidal benzotriazoles



Outcome

DSS – regularly updated for new compound classes:

- New testing strategies
- New testing data
- New models
- Actual integrated evaluations, including uncertainty and variability
- On-line and stand-alone tool



Operationalization of Goals

- Collection of basic information
- Data gap analysis + validation existing models/methods
- Generation new data and models
- Integration data/models in probabilistic risk assessment framework
- Carry out hazard assessment – quantify uncertainty (exp. data, model predictions, fate assessment models, effect assessment models)
- Integrate in website



Collection of data and models

- Experimental data intrinsic hazards – Screening Initial Data Set Dossier (SIDS)
- Models – Screening Initial Data Set Dossier (SIDS)
- Generation new data essential for validation and proper hazard/risk assessment
- Database data/models: dissemination purposes



Development/validation QSAR models

- Evaluate performance
- Similarity analysis and multivariate ranking methods for identification of priority chemicals to orient the experimental testing
- Develop new QSARs where gaps are identified due to lack of existing models or due to models of insufficient quality
- Documentation of the performance of the (final) models selected and developed – QMRF Reporting



Integration of QSARs within hazard and risk assessment

- Integration in probabilistic risk assessment framework: characterize variability/ uncertainty, sensitivity analyses with regard to contributions in overall risk assessment framework, modelling of variability
- Evaluate ECETOC TRA screening risk assessment tool
- Evaluate methods and decision points for establishing scientific validity and applicability domains for QSAR models
- Explore possibilities for economic valuation of substitution of chemicals from within chemical classes
- Policy and management: provision of recommendations on a viable management strategy for optimized testing and in-silico modeling of hazardous organic substances



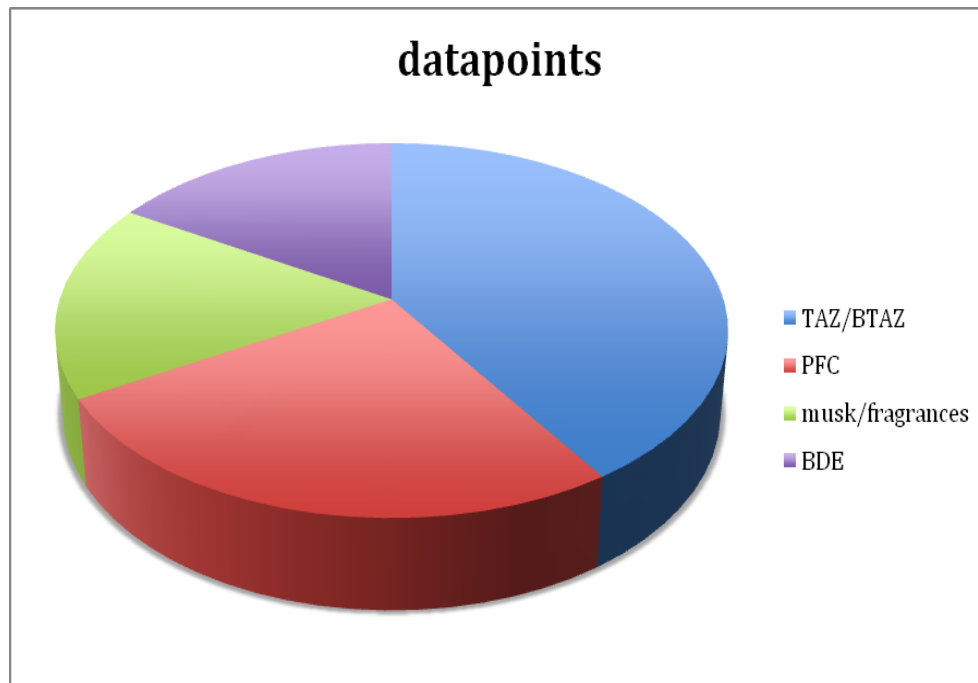
Outreach: website, newsletters/ workshops, stand-alone tools for dissemination of project results

- Development of on-line, stand-alone DSS: develop, publish, use QSAR/QSPR models for REACH
- Integration of the developed models with the QSAR Application Toolbox developed by OECD: establish the compatibility of the models with the (Q)SAR Model Reporting Format (QMRF) format
- Provision of a sustainable dissemination of project results by the WWW and as stand-alone tools
- Communication including newsletters and workshop(s)



SOME ACHIEVEMENTS (details during workshop)

- Data collection: 6838 exp. data – 2482 chemicals (dec. 2011)



- Lack of relevant (SIDS) data

Model collection



- Various models collected
- Various endpoints, various compound classes
- Local and general models: Local models generally higher accuracy
- Gap analysis as lead to further data collection and model development

- Majority of existing QSARs lack external validation;
- Definition Applicability Domain often lacking →
- Most models do not fulfill OECD principles QSAR validation

Model collection



Example: MLR-models
phys. Chem. Properties:

Endpoint	N _{obj}	Descriptors	R ²	Q ² _{LOO}	AD% _{243 BFR}
MP*	25	X2A	0.84	0.82	96%
LogVP*	34	T(O..Br)	0.99	0.98	83%
LogKoa*	30	T(O..Br)	0.97	0.97	82%
LogKow*	20	T(O..Br)	0.96	0.96	86%
Log H	7	BEHe7 Ext Val	not 0.97	0.93	56%
LogWS	12	Mor23m Ext Val	not 0.92	0.88	95%

Revision 6449 by itetko checked in on 2011-12-08 17:58:53. Built from null on 2011-12-08 18:10:38
 Welcome, Dear Miss.Kovarich!
 (1) My Account Logout

CADASTER

CAsE studies on the Development and Application of in-Silico Techniques for Environmental hazard and Risk assessment

Home Database Models A+ a-

Model applicator X Open the model profile X

Overview of the model

Review statistics of the model

Overview Applicability domain

Model name: LogWS_PFCs [rename], published in Prediction of Aqueous Solubility, Vapor Pressure and Critical Micelle Concentration for Aquatic Partitioning of Perfluorinated Chemicals (†). public identifier is 18
 Predicted property: Aqueous Solubility
 Training method: OLS
 modeled in log(mg/L)

[Dragon (blocks: 2 5)]
 Correl. limit: 1.0 Unique values: 0 Variance threshold: 0.0 Maximum value: 999999 UFS: false, [T(F..F), SIC1]

Uploaded model with external validation
 2 filtered descriptors
 Y = -0.419 -0.004*T(F..F) + 5.183*SIC1
 Calculated in 24 seconds
 Size: 6 Kb

Data Set	#	R2	q2	RMSE	MAE
Training set: PFC-LogWS Training set	20 records	0.76	0.76	0.84	0.62

Download model statistics in Excel format
 Create a copy of this model
 View configuration XML Export configuration XML
 Add QMRF url

RECALCULATE MODEL AND STATISTICS APPLY THE MODEL TO NEW COMPOUNDS

Summary models developed



Partner	Endpoint	Algorithm	Molecular Descriptors	Applicability Domain	Validation
PFCs					
RIVM	EC50 lettuce EC50 algae EC50 <i>D. magna</i> EC50 <i>C.sphaericus</i>	-Linear Regression based on nC (number of carbon atoms in the alkyl chain) -Interspecies Relationships <i>D. magna</i> - <i>C.sphaericus</i> and lettuce-algae	nC		R ²
B-TAZs					
UI	EC50 algae EC50 daphnia LC50 fish	-MLR regression (OLS) to all the endpoints -Interspecies Relationship Daphnia-Fish (prediction of EC50 fish)	1D-2D descriptors by Dragon (v. 5.5), PaDEL (open source) and CADASTER database	Leverage approach	Internal (R ² , Q ² _{LOO} , Q ² _{BOOT} , Y-sc etc)
	LC50 earthworm LD50 honeybee LD50 bird	Classification K-NN			External (Q2ext F1-F2-F3, R2m, CCC etc)
IVL	EC50 algae EC50 daphnia LC50 fish	PLS Regression	1D-2D-3D descriptors by Dragon software (v. 6.0)	DModX	Internal (R ² , Q ² , RMSEE) External (RMSEP)
LnU	LC50 fish	PLS Regression	descriptors by Dragon software (v. 6.0)		Internal (R ² , RMSE) External (Q2ext RMSE)
HMGU	EC50 algae EC50 daphnia LC50 fish	kNN, ASNN, FSMLR, PLS, MLRA, SVM	1D-2D-3D descriptors by CADASTER database		
IDEA	LC50 fish	MLR	Dragon 5.4 descriptors from SMILES	Leverage approach	Internal (R ² , Q ² _{LOO} , Q ² _{BOOT} , Y-sc etc) External (Q2ext F1-F2-F3, CCC etc)



Integration within Hazard/risk assessment

- Ongoing: development of QSAR integrated probabilistic RA platform
- Integration models, exp. data, exposure, effects in probabilistic
- Explicit characterisation/quantification of variability/uncertainty:
exp. data, predictive models
- Sensitivity analysis: SSD, fate model: 'classical' statistical methods +
Monte Carlo simulations



Some key findings

- No single “menu” Hazard Assessment 4 compound classes
- Sensitivity analysis: outcome dependent on fate and effect properties
- Key is: optimize hazard/risk assessment by integration tools + data, taking chemical diversity into account

Dissemination



Entry page QSPR-thesaurus database

Revision 2010-12-20 16:54:47 by 146.107.19.12 checked in on null. Built from null on null

Safari 5 on Mac - Not tested

Welcome, Guest! [Logout](#)



Case studies on the Development and Application of in-Silico Techniques for Environmental hazard and Risk assessment



[Home](#) ▾ [Database](#) ▾ [Models](#) ▾

A+ a-

Models applier browser

Step 1. Select a model from the list

Filter by model name: and property name: or by article id: Models visibility: [Public and private](#) ▾ [\[refresh\]](#)

1 - 10 of 10

MP_TAZs_remap , published by simona	predicts Melting Point using TAZ_MP_Training set (56)	MLRA	2010-11-12	CSV Excel Printer Empty
LogVP_TAZs_remap , published by simona	predicts Vapor Pressure using TAZ_LogVP_Training set (33)	MLRA	2010-11-12	CSV Excel Printer Empty
LogWS_TAZ , published by simona	predicts Aqueous Solubility using TAZ_LogWS_Training set (49)	MLRA	2010-11-11	CSV Excel Printer Empty
LogWS_PFCs , published by simona	predicts Aqueous Solubility using PFC_LogWS_Training set (20)	MLRA	2010-11-11	CSV Excel Printer Empty
LogVP_PFCs_remap(p4-2C) , published by simona	predicts Vapor Pressure using PFC_LogVP_Training set (35)	MLRA	2010-11-12	CSV Excel Printer Empty
Papa_PDE_Henry new , published by wolfram	predicts Henry using Ester_Papa_LogH (7)	MLRA	2010-07-05	CSV Excel Printer Empty
Papa_PDE_TM new , published by wolfram	predicts Melting Point using Ester_Papa_TM (25)	MLRA	2010-07-05	CSV Excel Printer Empty
Papa_PDE_Log_PL new , published by wolfram	predicts Vapor Pressure using Ester_Papa_Log_PL (34)	MLRA	2010-07-05	CSV Excel Printer Empty
Papa_PDE_Log_KOW new , published by wolfram	predicts logPow using Ester_Papa_LogKOW (20)	MLRA	2010-07-05	CSV Excel Printer Empty
Papa_PDE_LogKOA New , published by wolfram	predicts LogKoa using Ester_Papa_LogKOA (30)	MLRA	2010-07-05	CSV Excel Printer Empty

1 - 10 of 10

[Next>>](#)

Dissemination



Publically available models

Revision 2010-12-20 16:54:47 by 146.107.19.12 checked in on null. Built from null on null

Safari 5 on Mac - Not tested

Welcome, Guest! [Logout](#)



Case studies on the Development and Application of in-Silico Techniques for Environmental hazard and Risk assessment

[Home](#) [Database](#) [Models](#)

A+ a-

Models applier browser

Step 1. Select a model from the list

Filter by model name: and property name: or by article id: Models visibility: [Public and private](#) [\[refresh\]](#)

1 - 10 of 10

MP_TAZs_remap , published by simona	predicts Melting Point using TAZ_MP_Training set (56)	MLRA	2010-11-12	CSV Excel Print
LogVP_TAZs_remap , published by simona	predicts Vapor Pressure using TAZ_LogVP_Training set (33)	MLRA	2010-11-12	CSV Excel Print
LogWS_TAZ , published by simona	predicts Aqueous Solubility using TAZ_LogWS_Training set (49)	MLRA	2010-11-11	CSV Excel Print
LogWS_PFCs , published by simona	predicts Aqueous Solubility using PFC_LogWS_Training set (20)	MLRA	2010-11-11	CSV Excel Print
LogVP_PFCs_remap(p4-2C) , published by simona	predicts Vapor Pressure using PFC_LogVP_Training set (35)	MLRA	2010-11-12	CSV Excel Print
Papa_PDE_Henry new , published by wolfram	predicts Henry using Ester_Papa_LogH (7)	MLRA	2010-07-05	CSV Excel Print
Papa_PDE_TM new , published by wolfram	predicts Melting Point using Ester_Papa_TM (25)	MLRA	2010-07-05	CSV Excel Print
Papa_PDE_Log_PL new , published by wolfram	predicts Vapor Pressure using Ester_Papa_Log_PL (34)	MLRA	2010-07-05	CSV Excel Print
Papa_PDE_Log_KOW new , published by wolfram	predicts logPow using Ester_Papa_LogKOW (20)	MLRA	2010-07-05	CSV Excel Print
Papa_PDE_LogKOA New , published by wolfram	predicts LogKoa using Ester_Papa_LogKOA (30)	MLRA	2010-07-05	CSV Excel Print

1 - 10 of 10

[Next>>](#)

Dissemination



3D optimization within MOPAC/BOINS



Case studies on the Development and Application of in-Silico Techniques for Environmental hazard and Risk assessment

[Home](#) [About Us](#) [Cadaster](#) [Mopac](#) [Help](#)

Records   Export 

1 - 5 of 14 [≥](#) [>>](#)

Items per page:

Page of 3




ID:86911
Source:original


ID:86912
Source:balloon


ID:87780
Source:original



MolecularWeight: 500.127 ElectronNuclear: -16935.1328
LumoEnergy: -2.37 NuclearNuclearRepulsion: 52706.3348
TotalEnergy: -10441.57568 ResonanceEnergy: -477.3395
TotalElectrostaticInteraction: 175.6548 CoreRepulsion:
52706.33478 IonisationPotential: 11.97059 DipolSum: 2.383
ElectronElectron: 6957.8887 ElectronicEnergy: -
63147.91046 DipolHybrid: 1.564 HomoEnergy: -11.971
ElectronNuclearAttraction: -104396.3585
ElectronElectronRepulsion: 51865.6785 FinalHeat: -
889.86796 ExchangeEnergy: -162.6469
PrincipalMomentsInertiaC: 12121.329671
PrincipalMomentsInertiaA: 1993.578204
PrincipalMomentsInertiaB: 11603.289242 DipolPointCharge:
3.312

TotalEnergy: -10441.63752.....



23



Implementation QMRF form

www.CADASTER.eu

Home	REACH	CHALLENGE	Database	Partners	WP	Meetings	News	Forum	Login
------	-------	-----------	----------	----------	----	----------	------	-------	-------



CAsE studies on the Development and Application of in-Silico Techniques for Environmental hazard and Risk assessment

Home

QSAR Model Reporting Format

Displaying a model description

1. QSAR identifier

[-] Expand

1.1 QSAR identifier (title)

[Help](#)

Polybrominated Diphenyl Ethers, LogKOA

1.2 Other related models

[Help](#)

Polybrominated Diphenyl Ethers, LogKOW Polybrominated Diphenyl Ethers, Log_PL Polybrominated Diphenyl Ethers, TM Polybrominated Diphenyl Ethers, Henry

1.3 Software coding the model

[Help](#)

	Name:	URL:	Description:	Contact:
1	MOBY-DIGS (Ver. 1.0)	http://www.taletc.mi.it		

2. General information

[+] Expand

3. Defining the endpoint - OECD Principle 1

[+] Expand