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**

REACH

Minimised animal use:

1 – Use of validated *in silico* techniques: (Q)SAR/(Q)SPR

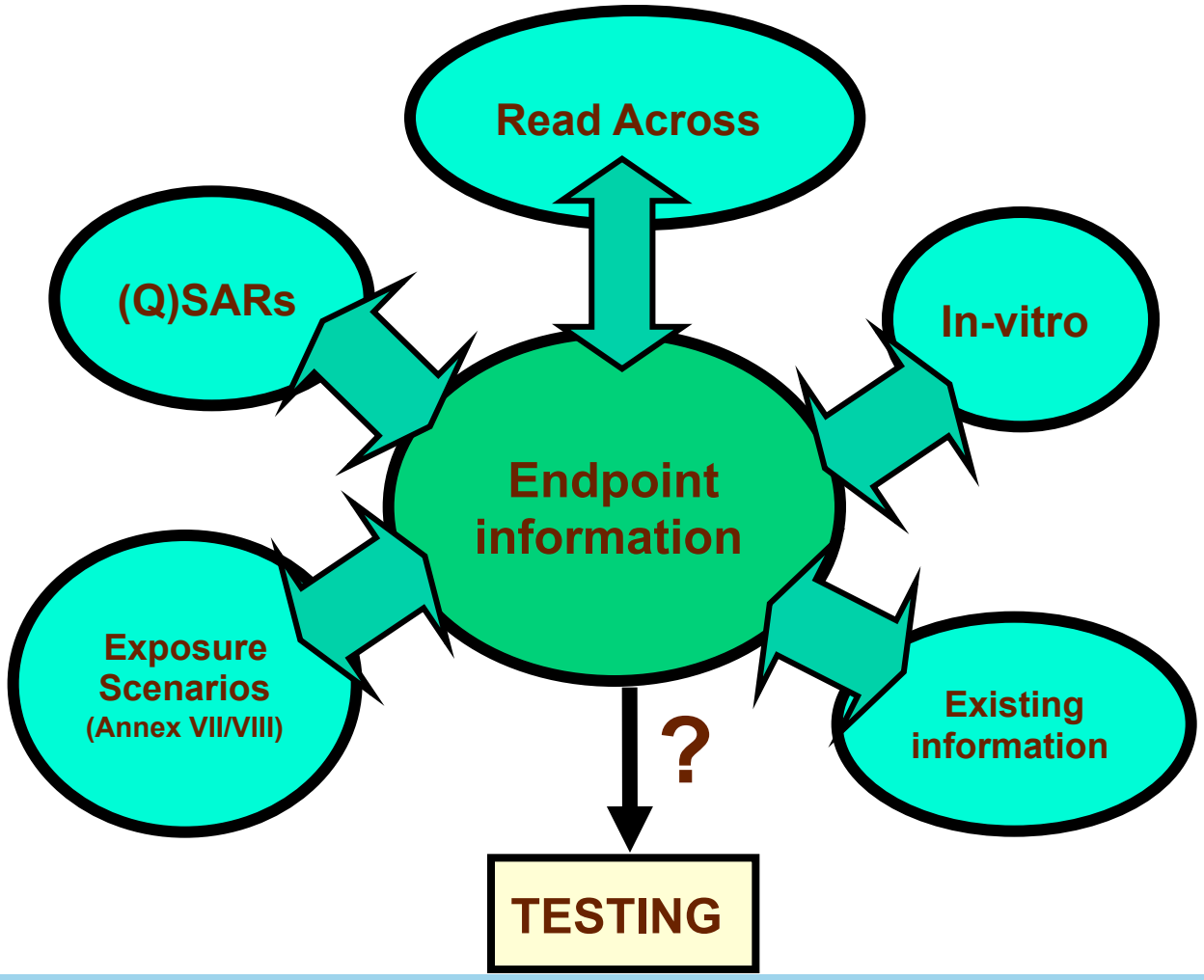
2 – New *in vitro* test methods, *in vivo* info analogues

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

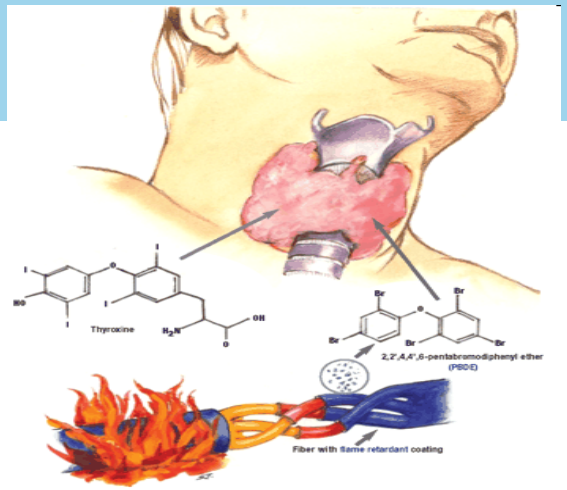
Provide full environmental hazard and risk assessment according to the REACH philosophy for chemicals belonging to 4 classes of emerging chemicals:

1 – Polybrominated diphenylethers (PBDE), typically class of **hydrophobic** chemicals that pose a threat to man and the environment.

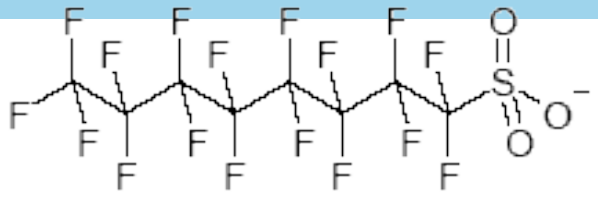
2 - Perfluoroalkylated substances and their transformation products, like perfluoroalkylated sulfonamides, alkanoic acids, sulfonates. **Persistent hydrophilic compounds** that may be **toxic** for man and environment.

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"

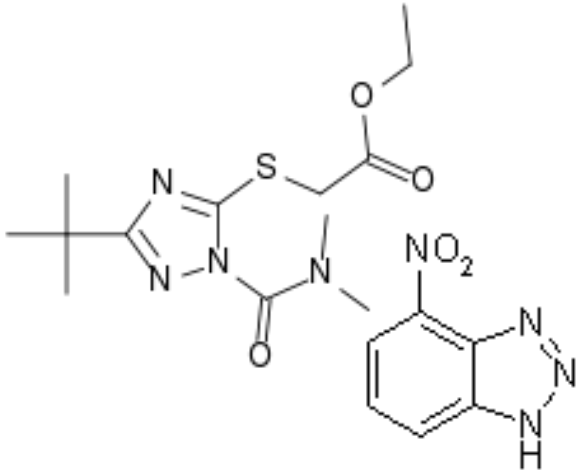


Classification of PFOS-compounds in 22 categories according to OECD

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

WP 1

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

WP 2

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

WP 3

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

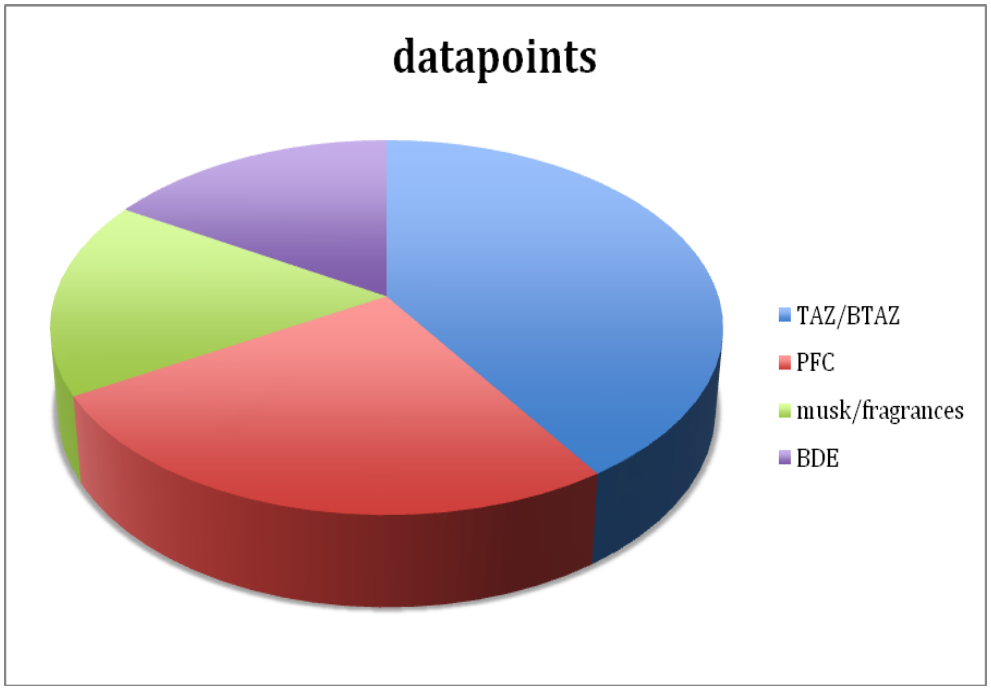
WP 4

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

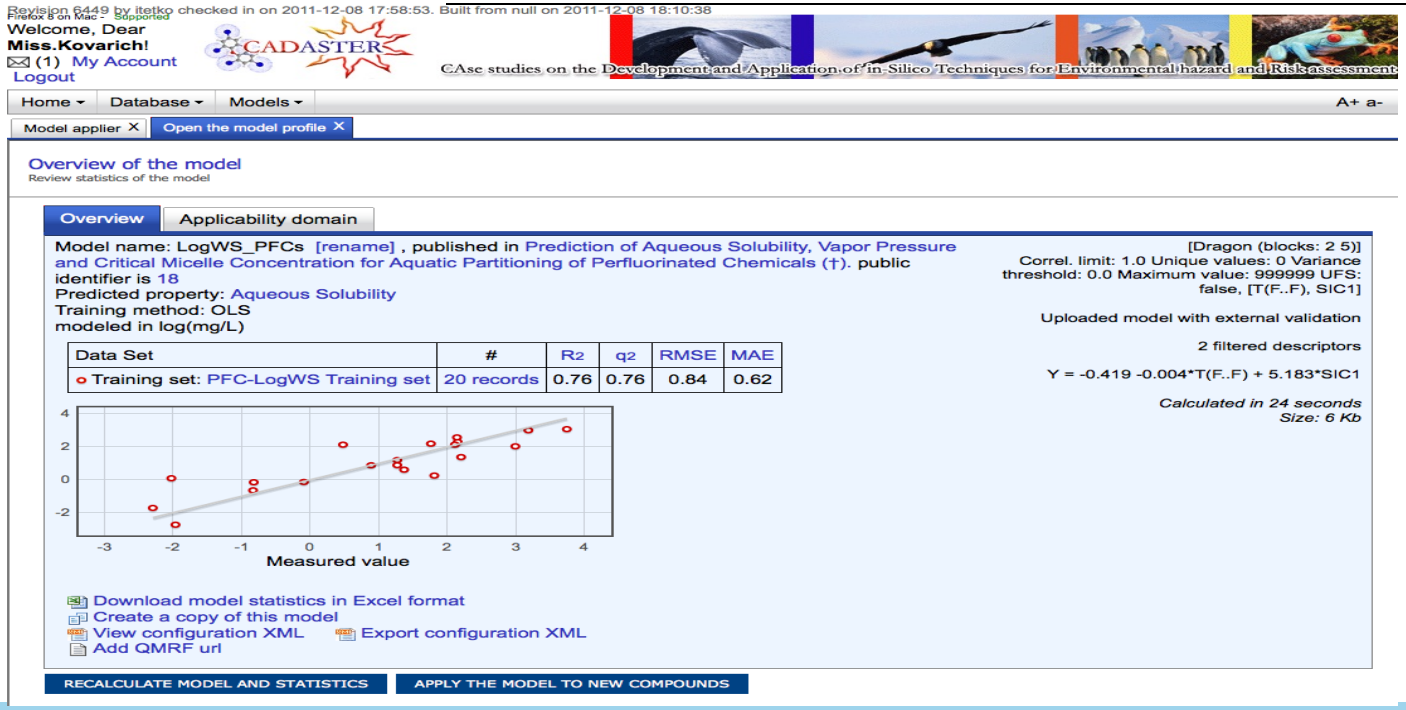
- Majority of existing QSARs lack external validation;
- Definition Applicability Domain often lacking →
- Most models do not fulfill OECD principles QSAR validation

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*	*not Ext Val*	0.97	0.93	56%
LogWS	12	*Mor23m*	*not Ext Val*	0.92	0.88	95%



Summary models developed

Partner	Endpoint	Algorithm	Molecular Descriptors	Applicability Domain	Validation
PFCs					
RIVM	EC50 lettuce EC50 algae EC50 *D. magna* EC50 *C.sphaericus*	-Linear Regression based on nC (number of carbon atoms in the alkyl chain) -Interspecies Relationships *D. magna-C.sphaericus* 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) External (Q2ext F1-F2-F3, R2m, CCC etc)
	LC50 earthworm LD50 honeybee LD50 bird	Classification K-NN			
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

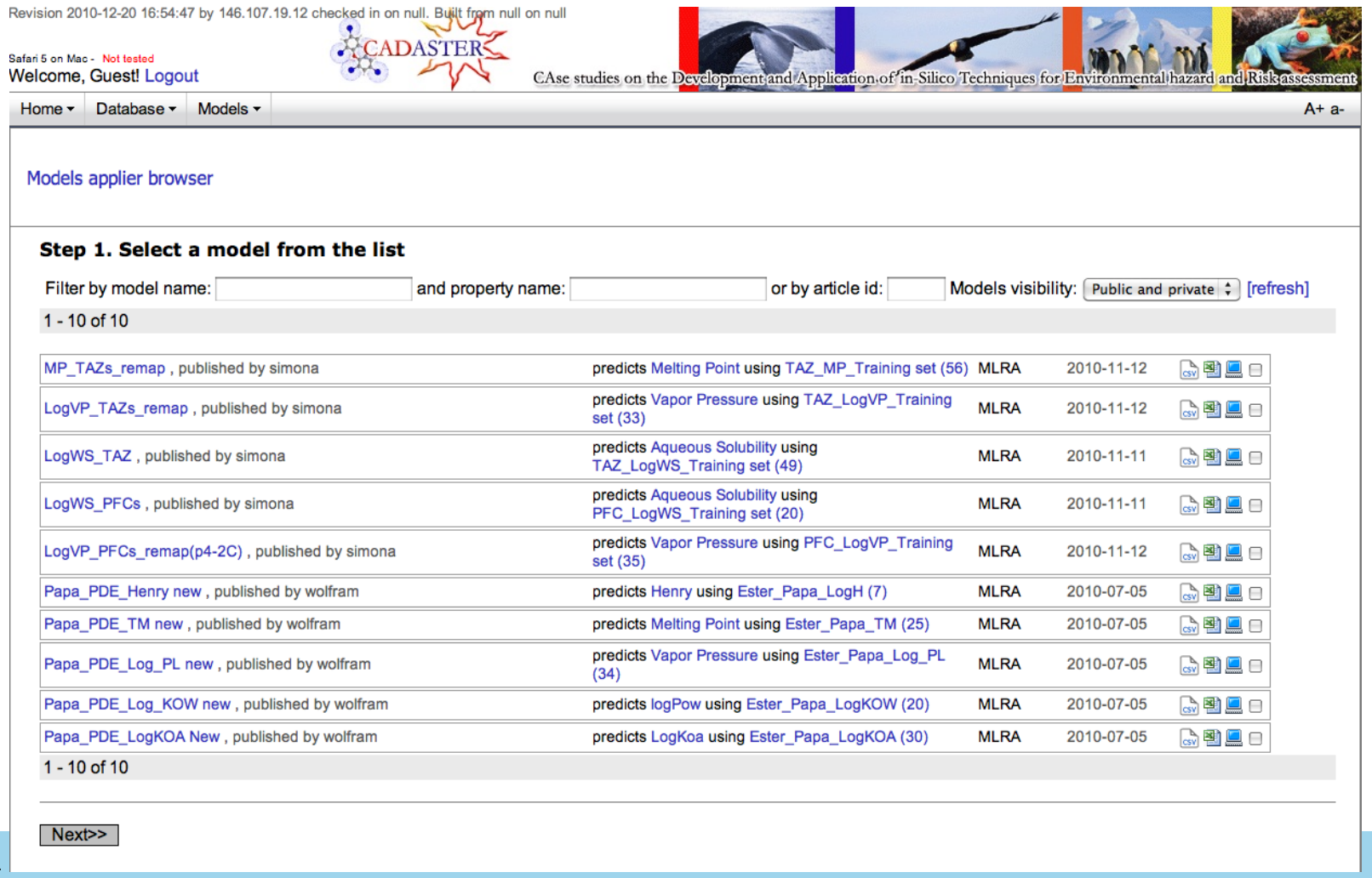
- Ongoing: development of QSAR integrated probabilistic RA platform
- Integration models, exp. data, exposure, effects in probabilistic
- Explicit characterisation/quantification of variability/uncertainty: exp. data, predictive models
- Sensitivity analysis: SSD, fate model: 'classical' statistical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

Some key findings

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

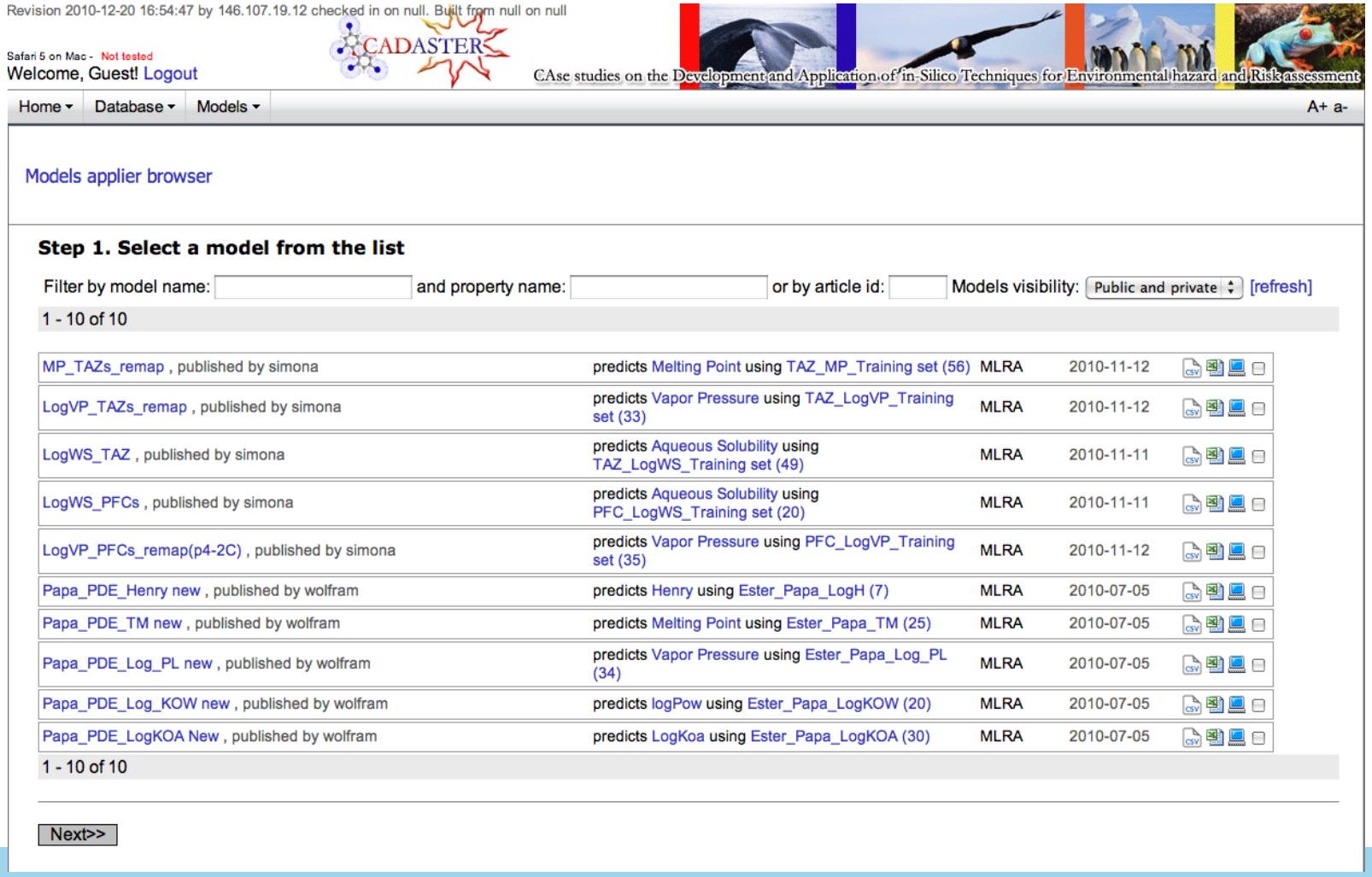
Dissemination

Entry page QSPR-thesaurus database



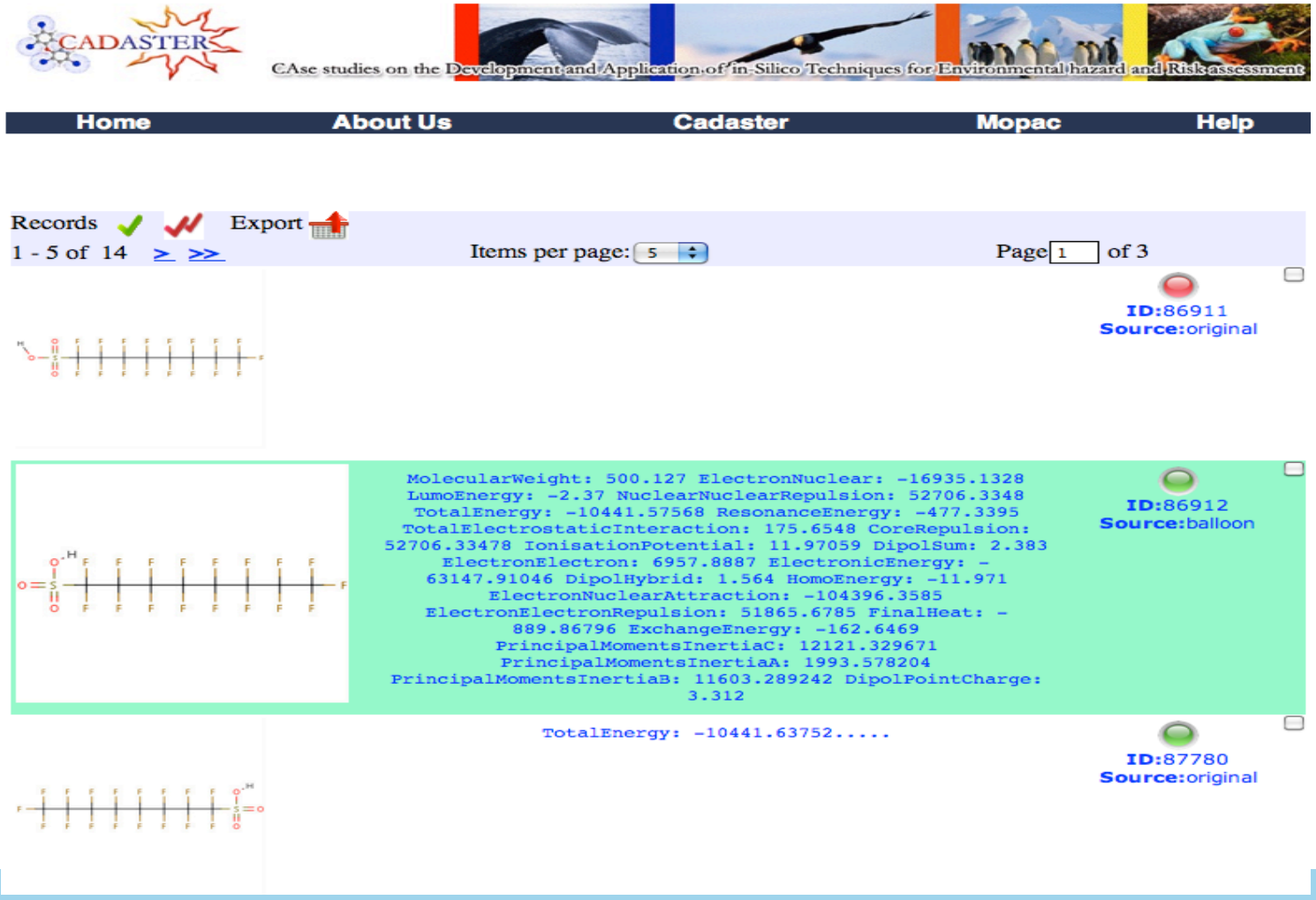
Dissemination

Publically available models



Dissemination

3D optimization within MOPAC/BOINS



Dissemination

Implementation QMRF form