Uso di QSAR e read across per predire proprietà di interesse tossicologico

EMILIO BENFENATI

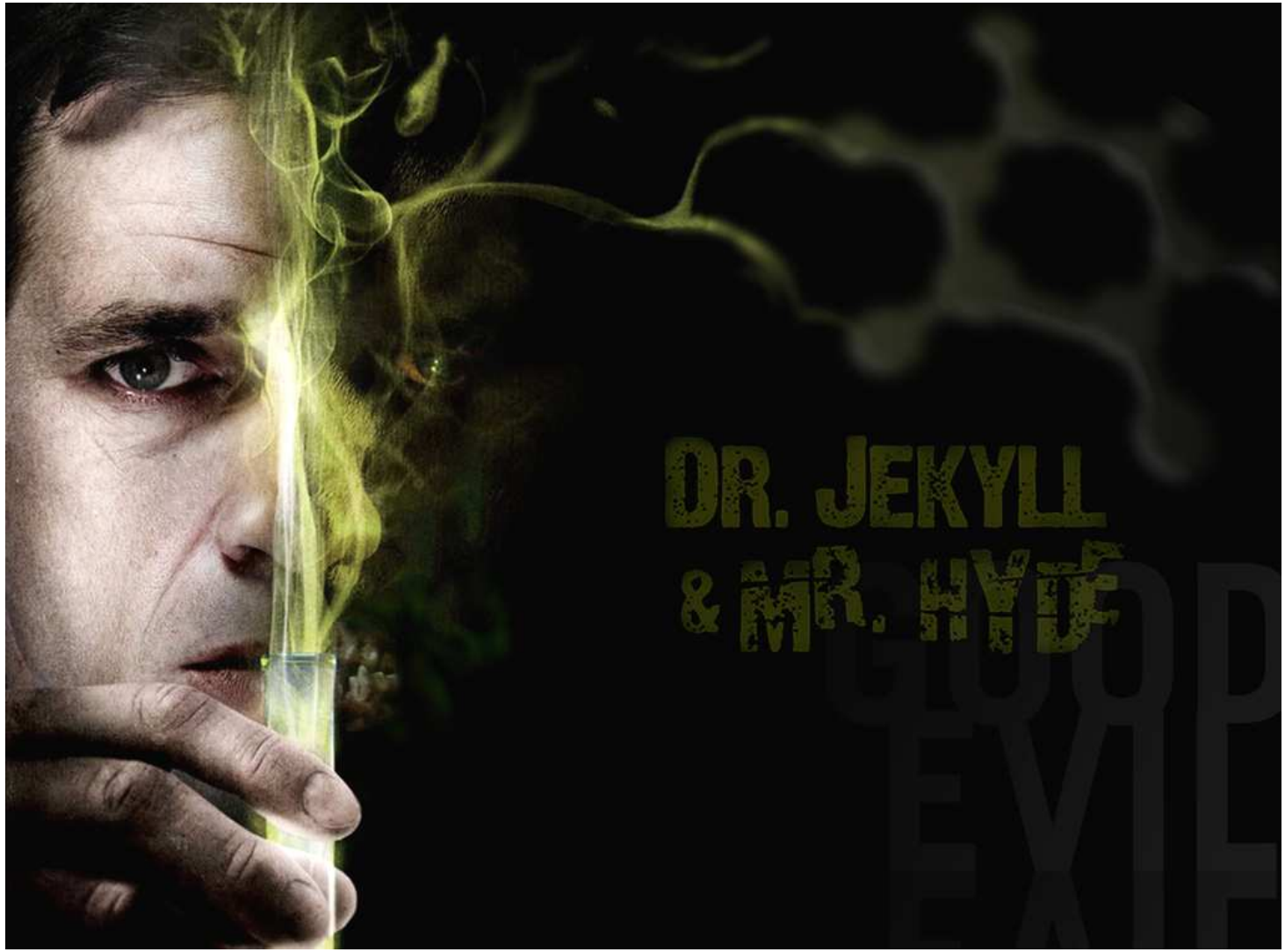
Istituto di Ricerche Farmacologiche Mario Negri

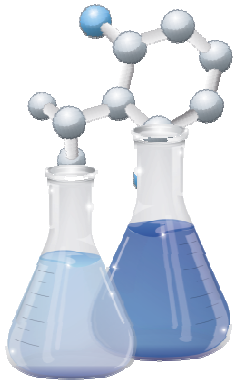
Laboratory of Environmental Chemistry and Toxicology

Corso teorico-pratico di valutazione della sicurezza dei cosmetici

SITOX UNIPRO, Milano,15-19 aprile, 2013

CHEMICALS: GOOD and EVIL

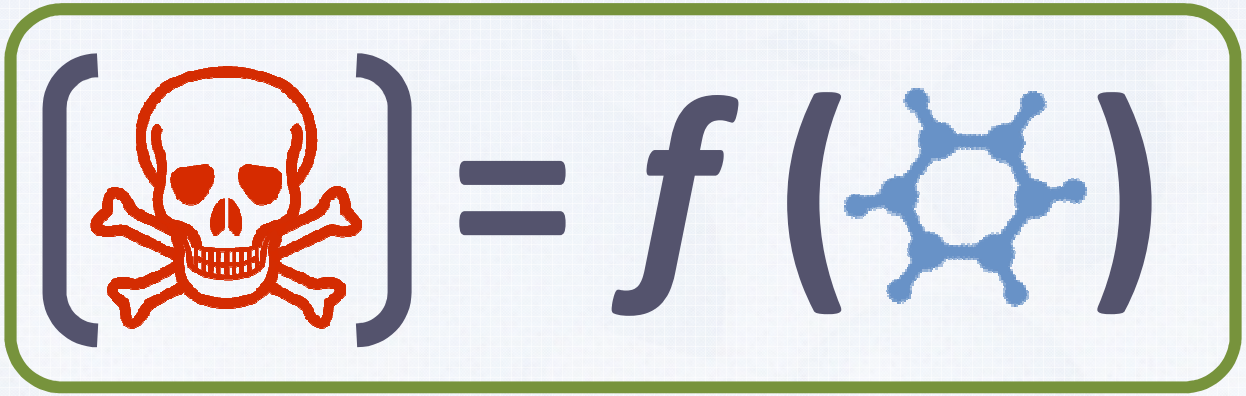




(Q)SAR

=

(Quantitative) Structure-Activity Relationship



IN SILICO


QSAR
0110011
01
0101

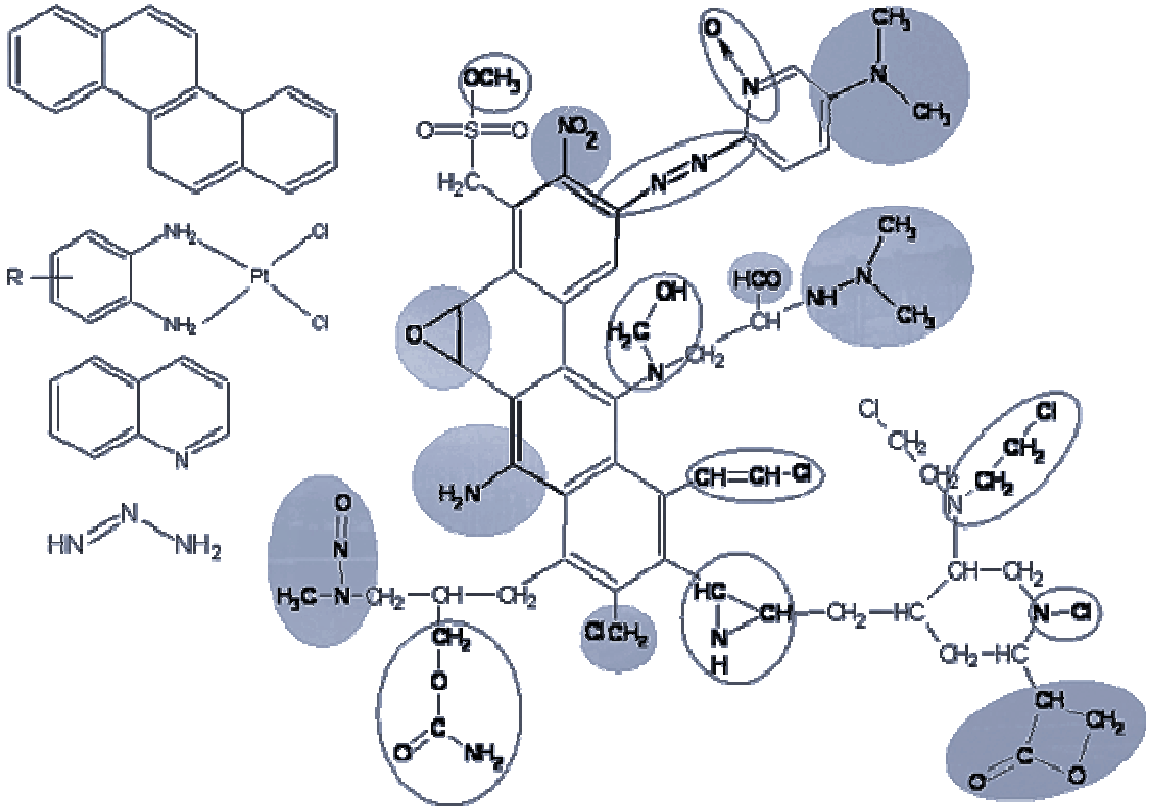

QSAR
0110011
01 0101

HUMAN EXPERTS have identified

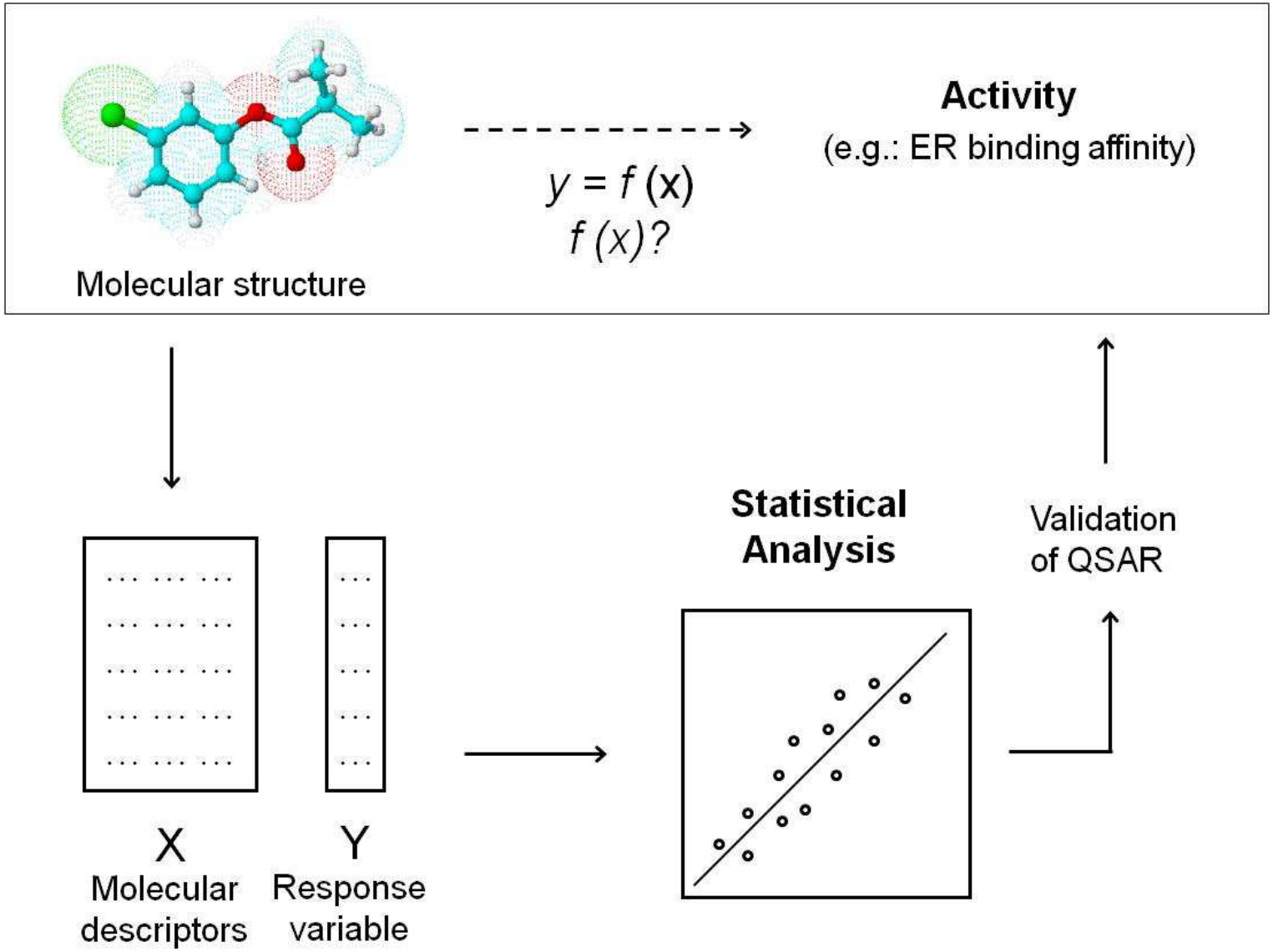
LINKS between

STRUCTURE and TOXICITY

ASHBY identified a list of RESIDUES for GENOTOXIC EFFECT



QSAR flow-chart



Simplified Molecular Input Line Entry Specification (SMILES™)

It was invented by Arthur and David Weininger founders of Daylight Chemical information Systems Inc.

Using ASCII strings for depicting chemical information!!

If ASCII strings are able to denote a feeling **:-)** , why not an organic formula?

O=[N+]([O-])c1c(c(c(c(c1C)[N+](=O)[O-])C(C)(C)C)[N+](=O)[O-])C

smiles:	62 bytes
MDL MOL:	2066 bytes
Connect Table:	998 bytes

Depicting Atoms

All atoms are depicted as their atomic symbols

C, N, O, P, S, F, Cl, Br, I

If they are not organic, or are acting with a non lowest normal valence they should go between brackets

[Fe], [S], [O-],...

Hydrogen should be removed unless is chemically meaningful

[H+], [C@@H], [OH-]

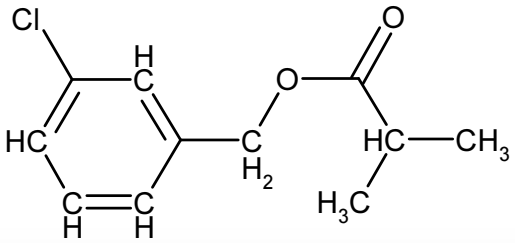
So:

C	Methane	CH_4
P	Phosphine	PH_3
O	Water	H_2O
Cl	Hydrochloric acid	HCl
[C]	Graphite/Diamond	C

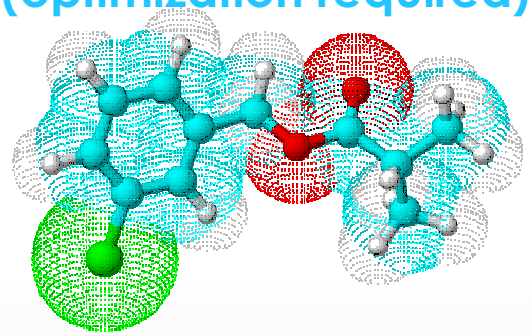
The procedure to

CALCULATE DESCRIPTORS

2D descriptors
(no optimization required)



3D descriptors
(optimization required)



The procedure adopted to calculate the 2D DESCRIPTORS may vary based on the different software requirement as *input file format*

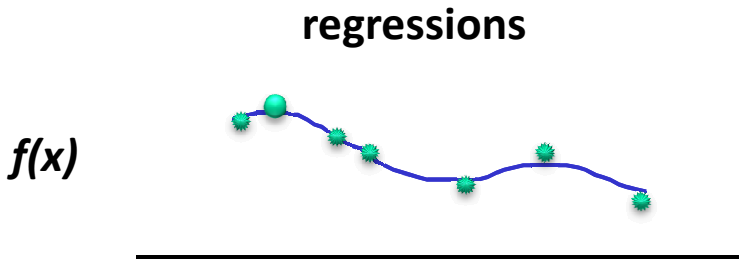
The 3D DESCRIPTORS are also affected by the *geometry optimization procedure*

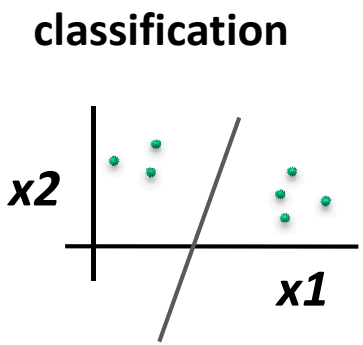
Many DESCRIPTORS FAMILIES:

- **Constitutional / information descriptors:** molecular weight, number of chemical elements, number of H-bonds or double bonds, …
- **Physicochemical descriptors:** lipophilicity, polarizability, …
- **Topological descriptors:** atomic branching and ramification
- **Electronic, geometrical and quantum-chemical descriptors**
- **Fragmental / structural keys** defining Booleans (bitmap) arrays

ALGORITHMS: CLASSIFIERS

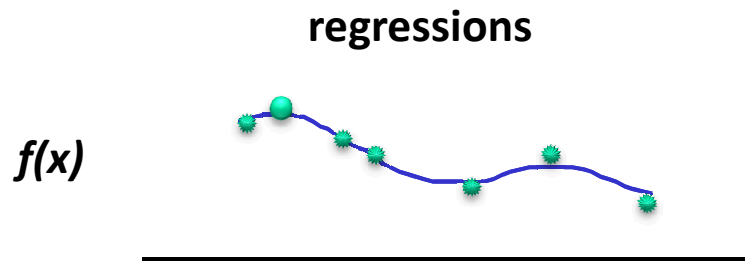
- Discriminant Analysis
- CART
- KNN
- Fuzzy logic
- Bayesian
- Self Organizing Map (SOM)
- Support Vector Machine (SVM)

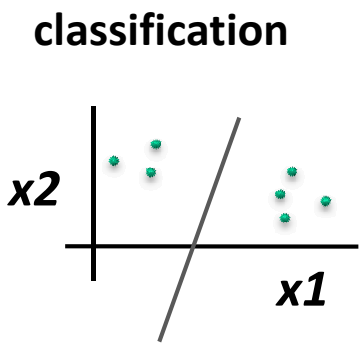




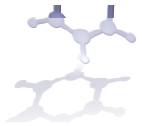
ALGORITHMS: REGRESSIONS

- Multi Variate Analysis (MVA)
- Partial Least Squares (PLS)
- Neural Networks (NN)
- Other algorithms
  (PCA, Genetic Algorithms)





MODEL PERFORMANCE



Robustness (training set)
- Leave-one-out, leave-some-out, bootstrap, etc.
- Y-scrambling

Prediction ability
- Prediction on an external set (TEST SET)
- False positives and false negatives

Applicability domain

Chemical and response space where the model can be applied

Statistical parameters - QSAR

Root-mean square error (RMSE):

average difference between the N predicted (A) and experimental (A') values

$$RMSE = \sqrt{\frac{\sum_{i=1}^{N} (A'(i) - A(i))^2}{N}}$$

Correlation coefficient (R^2):

degree of correlation between predicted (A) and experimental (A') values

$$R^2 = \frac{(\sum_{i=1}^{N} (A(i) - M_A)(A'(i) - M_{A'}))^2}{\sum_{i=1}^{N} (A(i) - M_A)^2 \sum_{i=1}^{N} (A'(i) - M_{A'})^2}$$

Fisher test (F):

it determines if the correlation is significant for at least x% compounds

$$F = \frac{\sum_{i=1}^{N} (A(i) - M_A)^2}{\sum_{i=1}^{N} (A'(i) - M_{A'})^2}$$

PRESS/SSY:

fraction of unexplained variance over the total variance

$$\frac{PRESS}{SSY} = \frac{\sum_{i=1}^{N} (A'(i) - A(i))^2}{\sum_{i=1}^{N} (A(i))^2}$$

STATISTICAL PARAMETERS: CLASSIFIER

CONFUSION MATRIX

ex., two classes (“positive” and “negative”) discrimination

		predicted: negative	predicted: positive
experimental: negative		TN	FP
experimental: positive		FN	TP

Accuracy (AC)

ratio of the total number of predictions that are correct

TN + TP / all

Sensitivity (Se)

ratio of positive predicted cases that are correct. High values preferred

TP / TP + FN

Specificity (Sp)

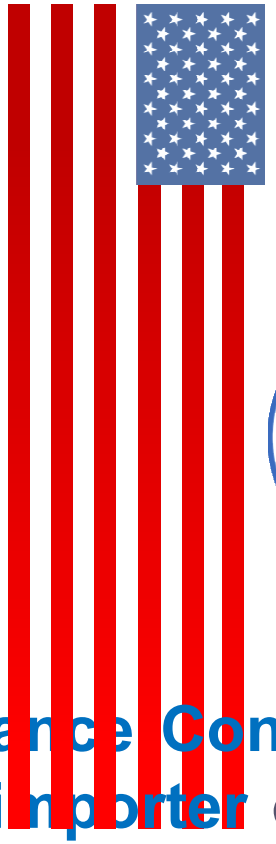
ratio of negative predicted cases that are correct

TN / TN + FP

False positives & negatives (FP, FN)

ratio (%) of positive and negative cases that are incorrectly classified.

united states / 1

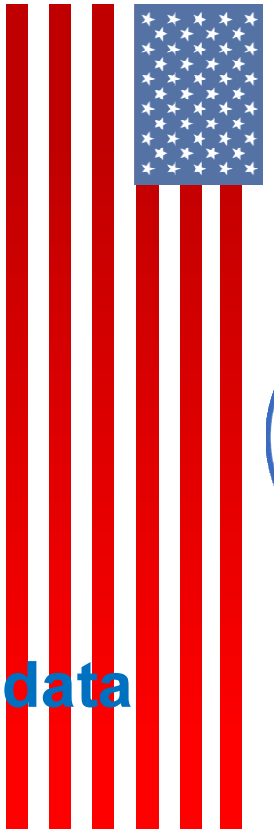


US epa

New Chemicals Program
Industrial Chemicals

Section 5 of TSCA (Toxic Substance Control Act) requires a manufacturer and/or importer of a new chemical substance to submit a premanufacture notice (PMN) to US EPA *90 days before* commencing manufacture or import of the new chemical

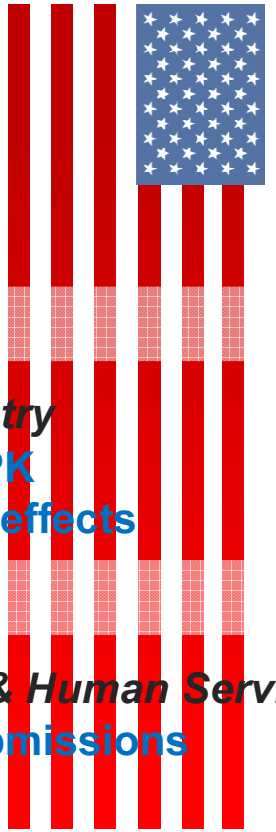
united states / 2



US epa

- Decisions often made in the **absence** of any **experimental data**
- **SAR methods** and **(Q)SAR developed** to help **reviews**
- US EPA evaluates approximately **1500-2000 PMN cases** a year

united states / 3



ATSDR

Agency for Toxic Substances and Disease Registry

- Toxicity prediction - QSARs based on PBPK
- Benchmark Dose (BMD) for human health effects

FDA

Food and Drug Administration - Dept. of Health & Human Services

- Carcinogenicity - data from regulatory submissions used to develop MULTICASE

NTP

National Toxicology Program

- Carcinogenicity - tested commercial software

NIOSH

National Institute for Occupational Safety and Health

- Use of SARs for hazard alerts for *Current Intelligence Bulletins*

REACH and Cosmetic regulation

The REACH APPROACH

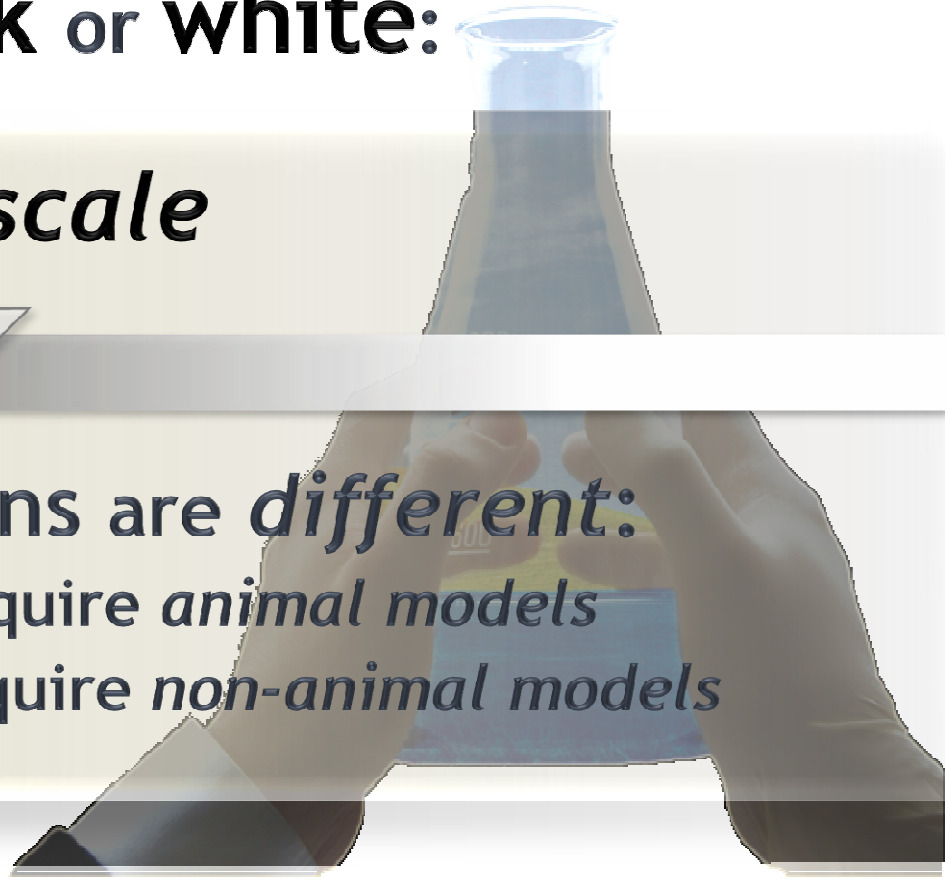
is not black or white:

- there is a *grey scale*

- other legislations are *different*:

  e.g. pesticides require *animal models*

  cosmetics require *non-animal models*



SEVEN REASONS to use QSAR

- **Innovation** *(also in view of milions of new data - ToxCast)*
- **Time for experiments**
- **Occurrence of enough laboratories/resources**
- **Reduction of costs**
- **Use of animals**
- **Prioritization needs**
- **Pro-active approach for “greener” chemicals**

The AIM of the REACH REGULATION

Article 1 : AIM and SCOPE

The purpose of this Regulation is to ensure a high level of protection of human health and the environment, *including the promotion of alternative methods for assessment of hazards of substances*, as well as the free circulation of substances on the internal market while enhancing competitiveness and innovation.

REACH AND QSAR

AIM and STRATEGY

REACH TARGET is ***MAN*** and ***ENVIRONMENT***


TARGET

IN VIVO models

IN SILICO models

REACH AND QSAR

According to REACH Regulation (Annex XI)

a QSAR Model is VALID IF

1. the model is recognized *scientifically valid*;
2. the substance is included in the *applicability domain* of the model;
3. results are adequate for *classification and labelling* and for *risk assessment*;
4. adequate *documentation of the methods* is provided.

ANTARES

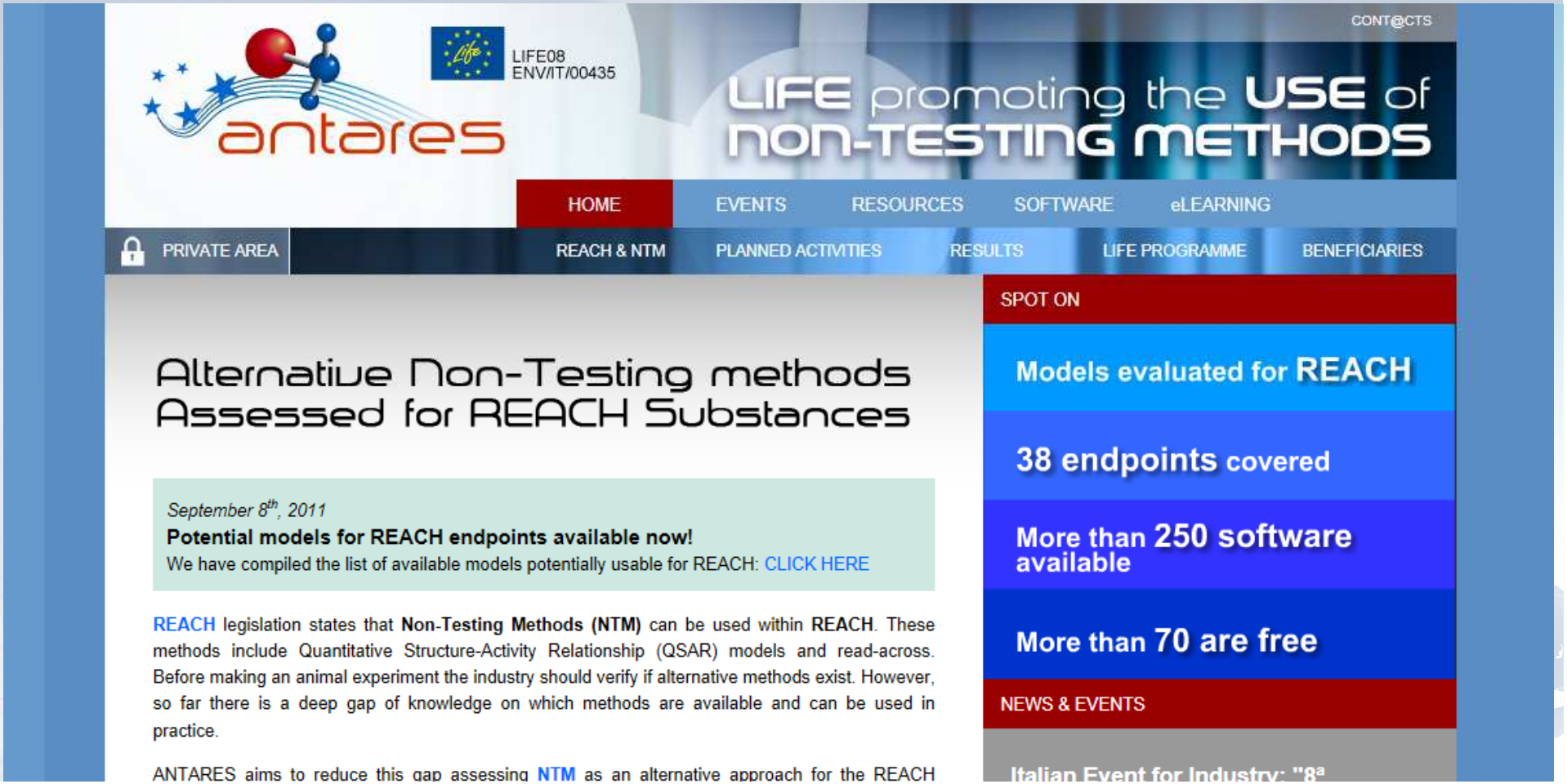
Evaluating the existence and suitability of Non-Testing Methods for REACH

Alternative Non-Testing methods Assessed for REACH Substances

Contract LIFE08 ENV/IT/000435

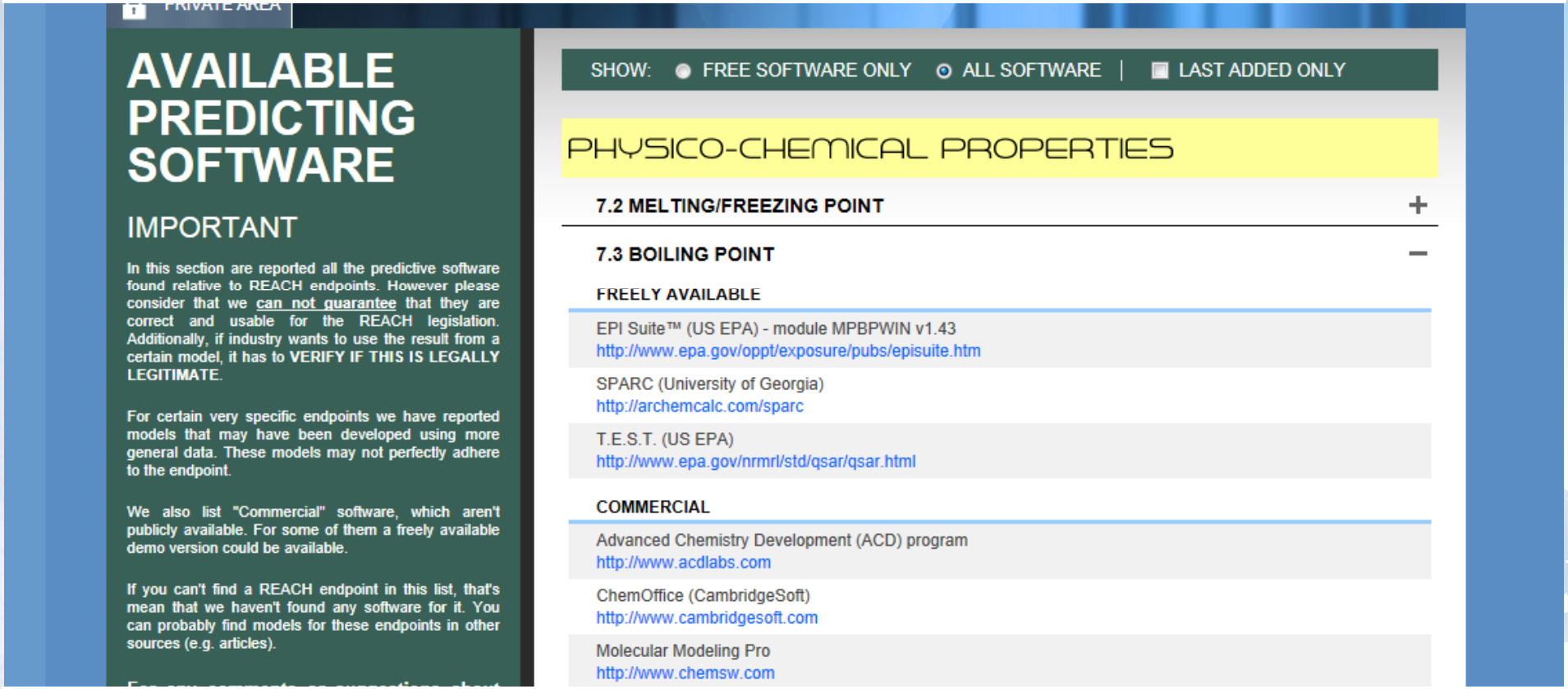
www.antares-life.eu

ANTARES


antares
LIFE08
ENV/IT/00435
CONT@CTS
LIFE promoting the USE of NON-TESTING METHODS
HOME
EVENTS
RESOURCES
SOFTWARE
eLEARNING
PRIVATE AREA
REACH & NTM
PLANNED ACTIVITIES
RESULTS
LIFE PROGRAMME
BENEFICIARIES
Alternative Non-Testing methods Assessed for REACH Substances
September 8th, 2011
Potential models for REACH endpoints available now!
We have compiled the list of available models potentially usable for REACH: CLICK HERE
REACH legislation states that Non-Testing Methods (NTM) can be used within REACH. These methods include Quantitative Structure-Activity Relationship (QSAR) models and read-across. Before making an animal experiment the industry should verify if alternative methods exist. However, so far there is a deep gap of knowledge on which methods are available and can be used in practice.
ANTARES aims to reduce this gap assessing NTM as an alternative approach for the REACH
SPOT ON
Models evaluated for REACH
38 endpoints covered
More than 250 software available
More than 70 are free
NEWS & EVENTS
Italian Event for Industry: "8ª

ANTARES

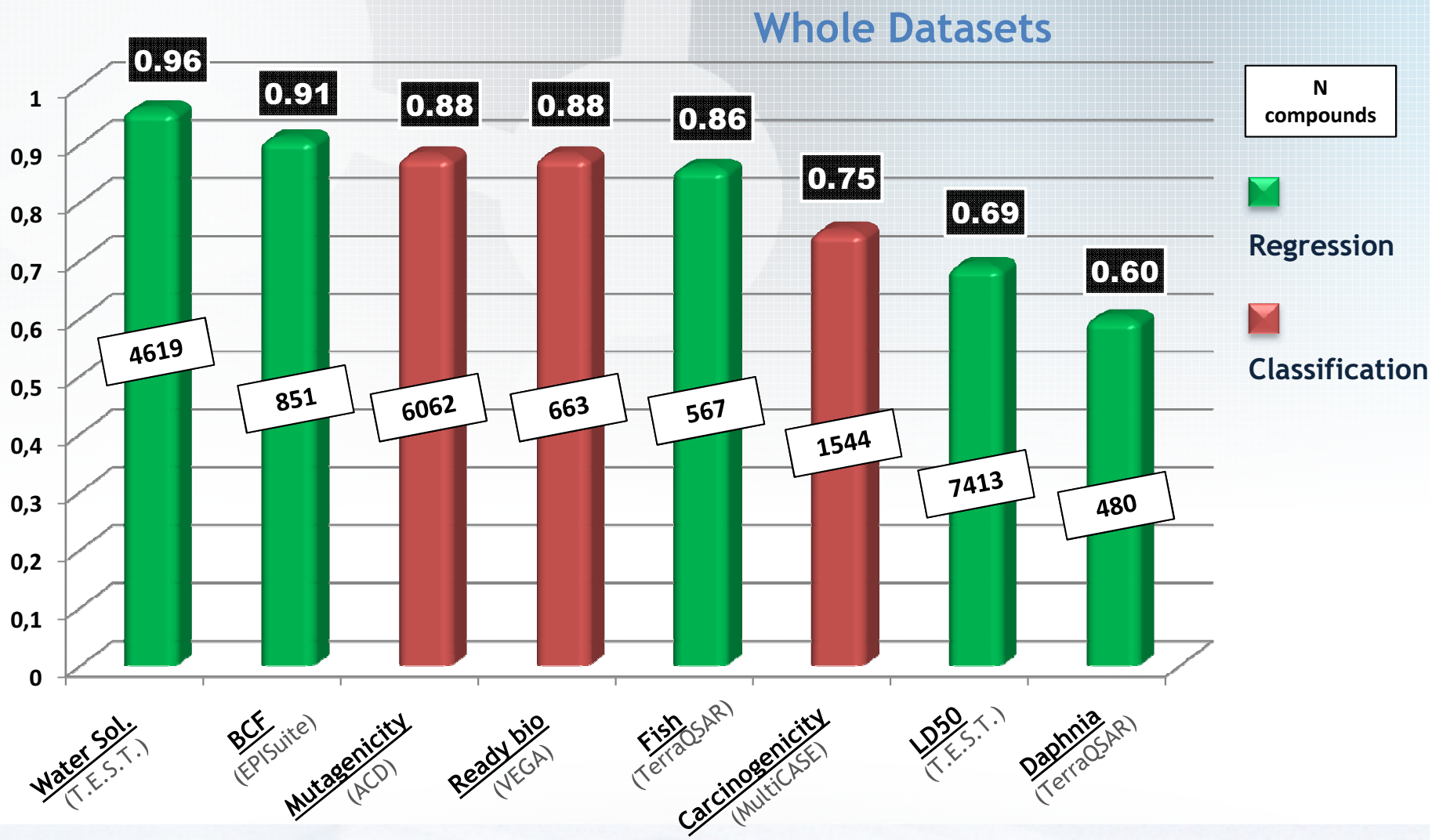

CONT@CTS
antares
LIFE08
ENV/IT/00435
LIFE promoting the USE of NON-TESTING METHODS
HOME
EVENTS
RESOURCES
SOFTWARE
eLEARNING
PRIVATE AREA
AVAILABLE PREDICTING SOFTWARE
IMPORTANT
In this section are reported all the predictive software found relative to REACH endpoints. However please consider that we can not guarantee that they are correct and usable for the REACH legislation. Additionally, if industry wants to use the result from a certain model, it has to VERIFY IF THIS IS LEGALLY LEGITIMATE.
For certain very specific endpoints we have reported models that may have been developed using more general data. These models may not perfectly adhere to the endpoint.
We also list "Commercial" software, which aren't publicly available. For some of them a freely available
SHOW: FREE SOFTWARE ONLY ALL SOFTWARE LAST ADDED ONLY
PHYSICO-CHEMICAL PROPERTIES
7.2 MELTING/FREEZING POINT
7.3 BOILING POINT
7.4 RELATIVE DENSITY
7.5 VAPOUR PRESSURE
7.6 SURFACE TENSION
7.7 WATER SOLUBILITY
7.8 PARTITION COEFFICIENT n-Octanol/Water


PRIVATE AREA
AVAILABLE PREDICTING SOFTWARE
IMPORTANT
In this section are reported all the predictive software found relative to REACH endpoints. However please consider that we can not guarantee that they are correct and usable for the REACH legislation. Additionally, if industry wants to use the result from a certain model, it has to VERIFY IF THIS IS LEGALLY LEGITIMATE.
For certain very specific endpoints we have reported models that may have been developed using more general data. These models may not perfectly adhere to the endpoint.
We also list "Commercial" software, which aren't publicly available. For some of them a freely available demo version could be available.
If you can't find a REACH endpoint in this list, that's mean that we haven't found any software for it. You can probably find models for these endpoints in other sources (e.g. articles).
SHOW: FREE SOFTWARE ONLY ALL SOFTWARE LAST ADDED ONLY
PHYSICO-CHEMICAL PROPERTIES
7.2 MELTING/FREEZING POINT +
7.3 BOILING POINT −
FREELY AVAILABLE
EPI Suite™ (US EPA) - module MPBPWIN v1.43
http://www.epa.gov/oppt/exposure/pubs/episuite.htm
SPARC (University of Georgia)
http://archemcalc.com/sparc
T.E.S.T. (US EPA)
http://www.epa.gov/nrmrl/std/qsar/qsar.html
COMMERCIAL
Advanced Chemistry Development (ACD) program
http://www.acdlabs.com
ChemOffice (CambridgeSoft)
http://www.cambridgesoft.com
Molecular Modeling Pro
http://www.chemsw.com

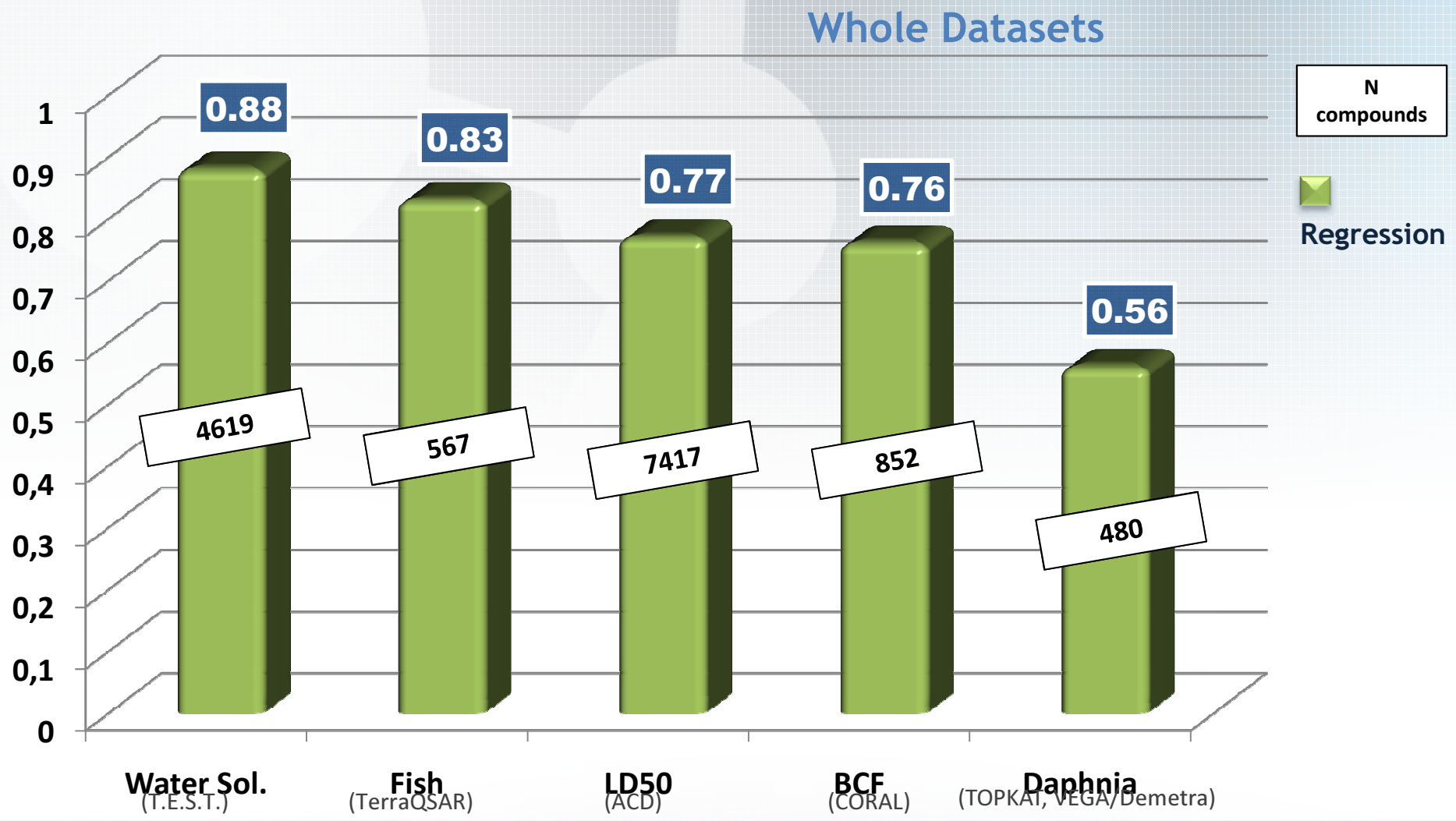
FOCUS ON 8 ENDPOINTS

Mutagenicity (Ames) Carcinogenicity LD50	HUMAN TOXICITY
Fish Acute Toxicity Daphnia Acute Toxicity	ECOTOXICOLOGY
BCF Ready Biodegradability	ENVIRONMENTAL
Water Solubility	PHYSICO-CHEMICAL

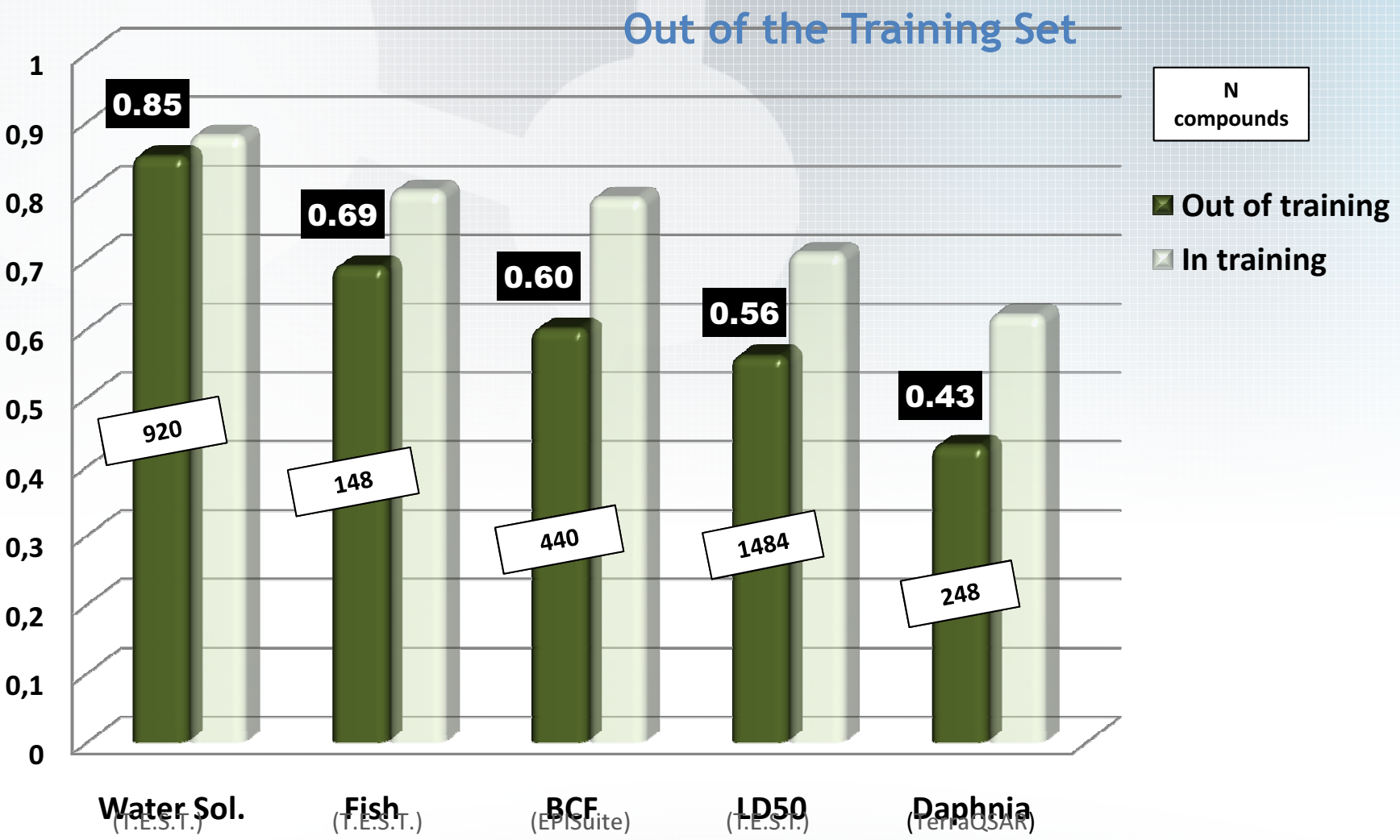
ACCURACY for the 8 endpoints



R^2 for 5 endpoints

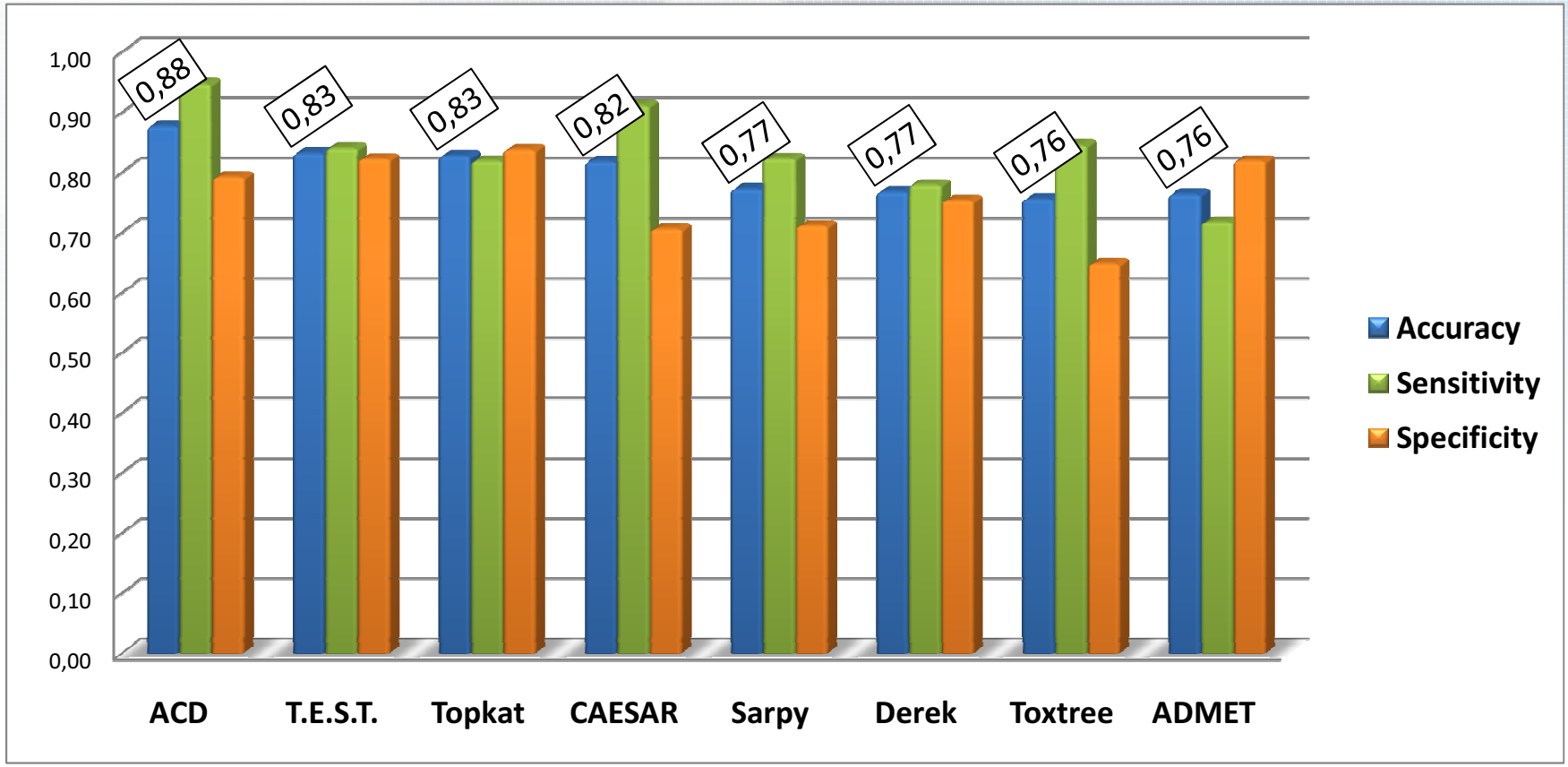


R^2 results considering new compounds



MUTAGENICITY: Performance

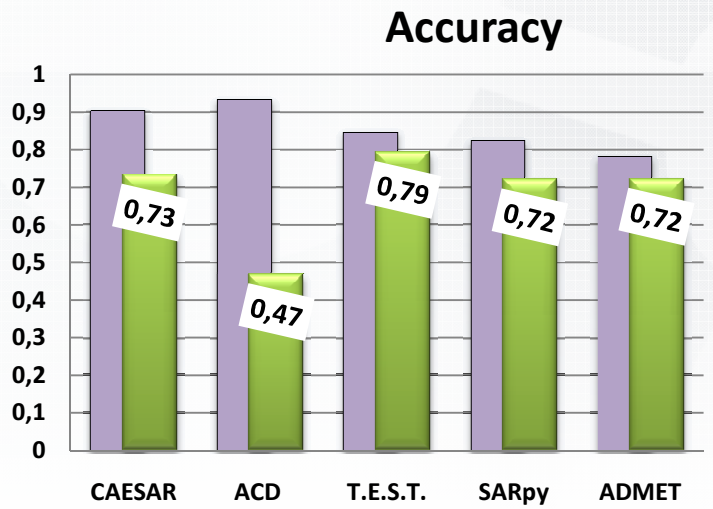
Total dataset (6065 compounds)



The first 4 models showed the best accuracy values very close to the in vitro reproducibility of Ames test (0.85)

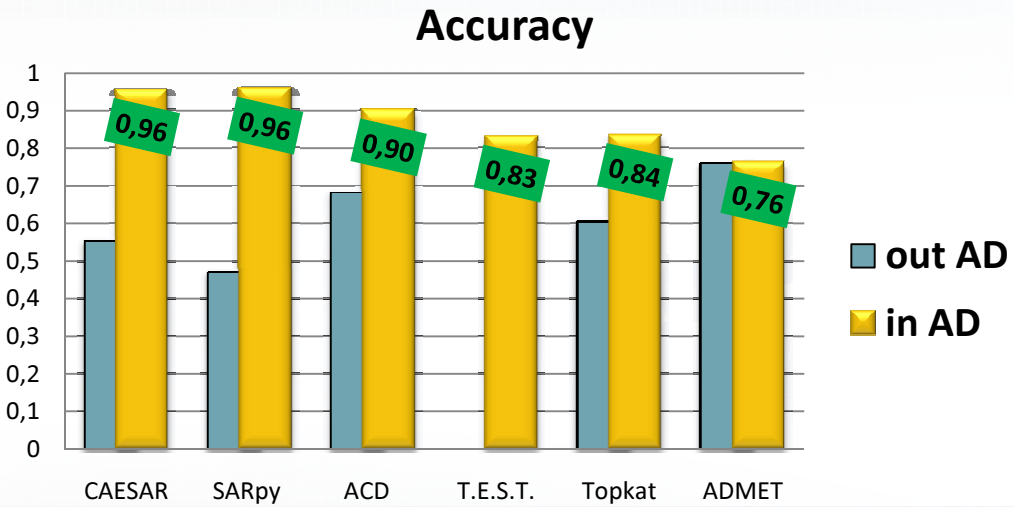
MUTAGENICITY: Performance

In & out train chemicals and in & out Applicability Domain



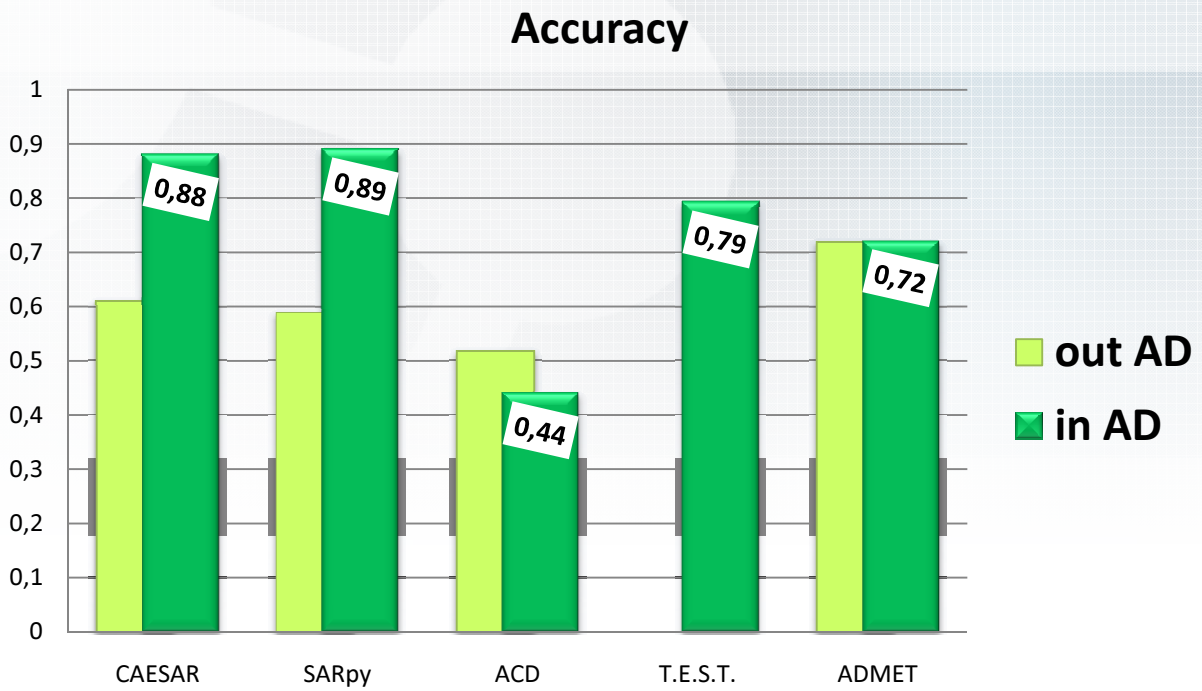
Excluding compounds in the training set: T.E.S.T. and CAESAR gave the highest accuracy. There is a decrease in the predictive performance considering molecules out of training set of the models.

An increase in the performance was seen after selecting the compounds inside the Applicability Domain for each model.



MUTAGENICITY: Performance

Chemicals out of train distributed by AD



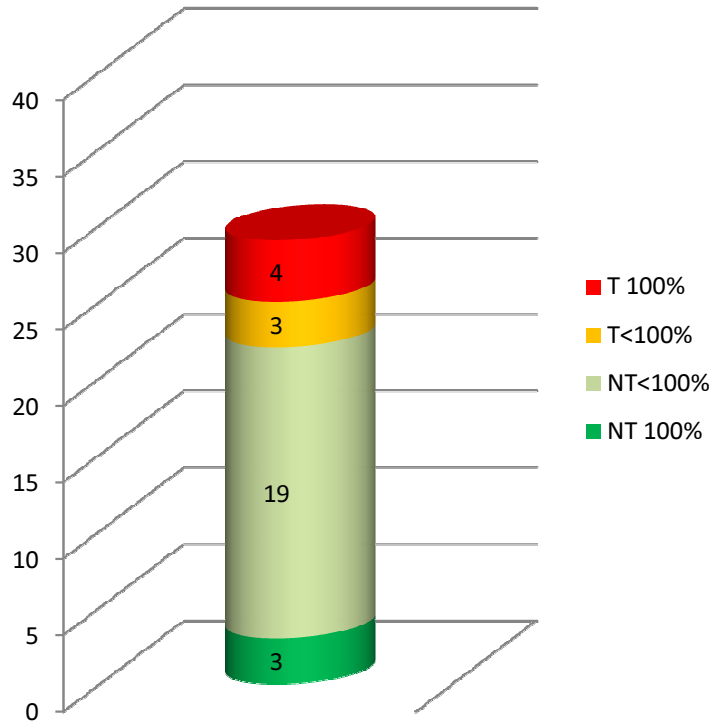
Applying the information on the applicability domain improves results.

For compounds out of training and within AD, CAESAR and SARpy gave the highest sensitivity.

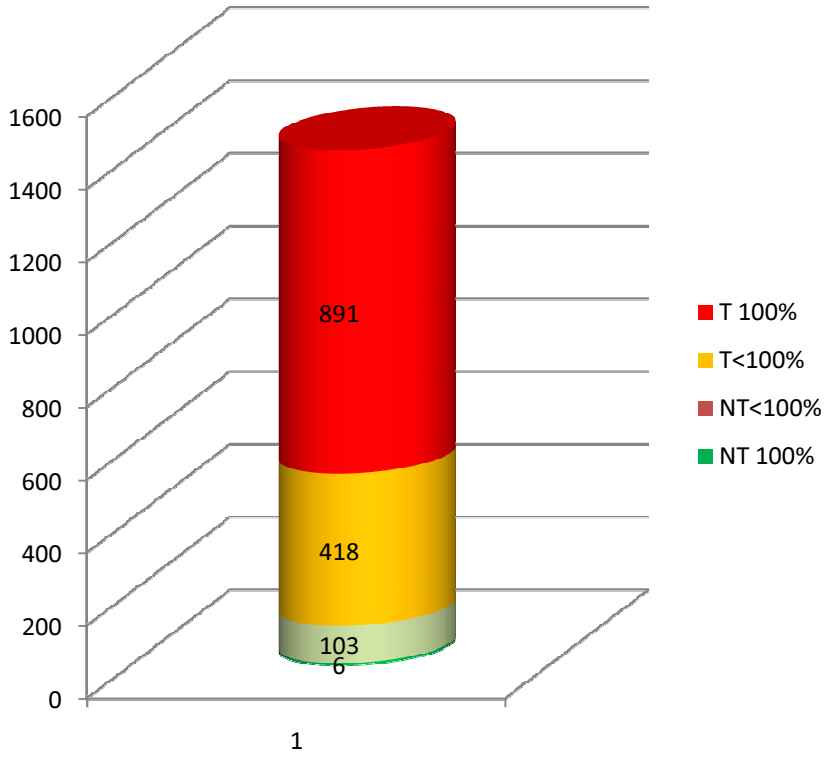
COMPARISON OF PREDICTION AND EXPERIMENTAL RESULTS:
False Negative and True Positive

	CHEMICALS	PREDICTION MODELS								EXPERIMENTAL TESTS RESULTS	
	CAS NUMBER	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**Hansen**	**Toolbox**
FNs common to all models	24280-93-1	NT	NT	NT	NT	NT	NT	NT	NT	T	NT
	538-23-8	NT	NT	NT	NT	NT	NT	NT	NT	T	T
	57166-92-4	NT	NT	NT	NT	NT	NT	NT	NT	T	T
	60129-60-4	NT	NT	NT	NT	NT	NT	NT	NT	T	T
	629-14-1	NT	NT	NT	NT	NT	NT	NT	NT	T	T
	68-26-8	NT	NT	NT	NT	NT	NT	NT	NT	T	T
	70-54-2	NT	NT	NT	NT	NT	NT	NT	NT	T	T
	7235-40-7	NT	NT	NT	NT	NT	NT	NT	NT	T	NT
	80-13-7	NT	NT	NT	NT	NT	NT	NT	NT	T	NT
	119-36-8	NT	NT	NT	NT	NT	NT	NT	NT	T	T
TPs common to all models	100-13-0	T	T	T	T	T	T	T	T	T	T
	100-16-3	T	T	T	T	T	T	T	T	T	T
	100-32-3	T	T	T	T	T	T	T	T	T	T
	100527-20-6	T	T	T	T	T	T	T	T	T	T
	100593-23-5	T	T	T	T	T	T	T	T	T	T
	10061-01-5	T	T	T	T	T	T	T	T	T	T
	100924-64-9	T	T	T	T	T	T	T	T	T	T
	101043-65-6	T	T	T	T	T	T	T	T	T	T
	10125-76-5	T	T	T	T	T	T	T	T	T	T
	10159-53-2	T	T	T	T	T	T	T	T	T	T

Spotting uncertain data



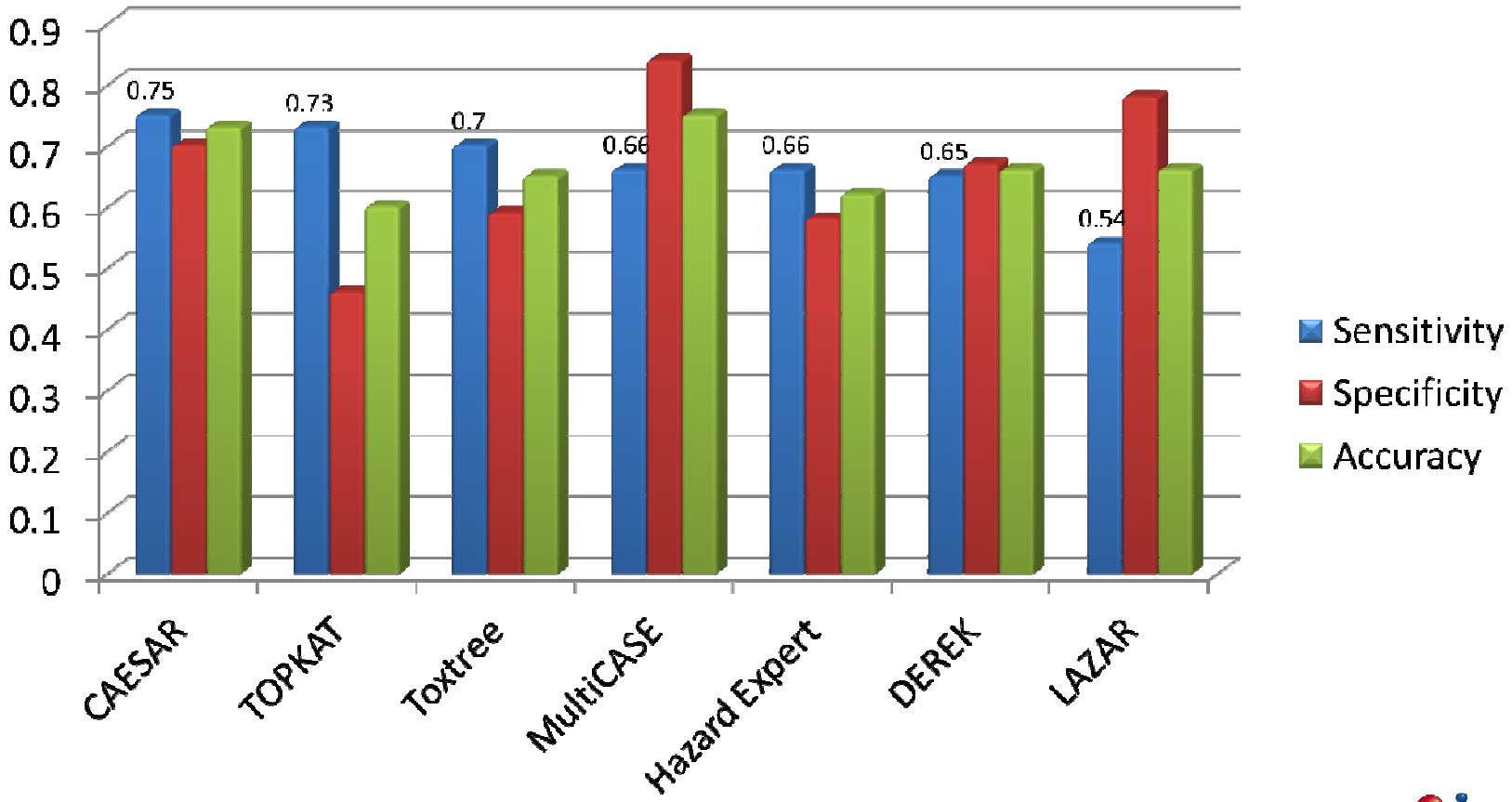
False Negatives



True Positives

Carcinogenicity: Results

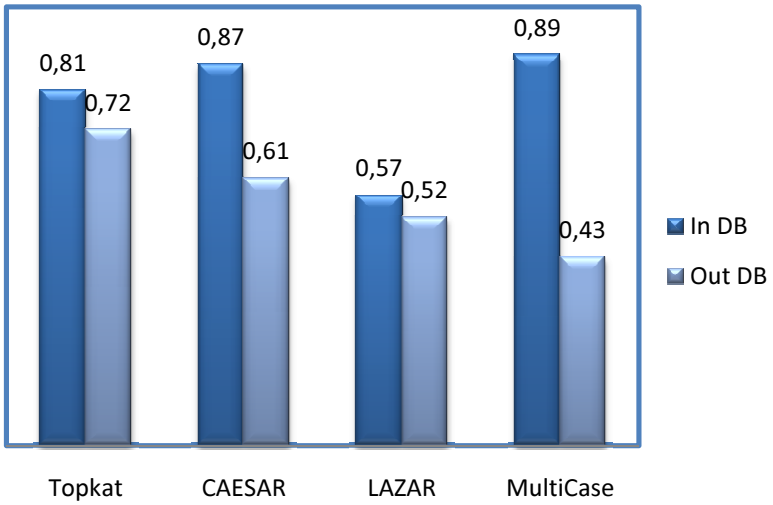
Predictions on the 1544 compounds (CPDB+Leadscope) of the seven programs



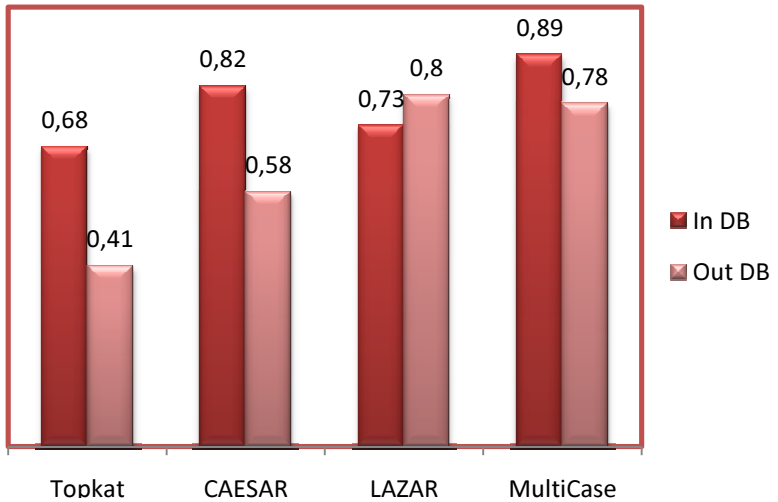
Carcinogenicity: Results for Training/out

Performance of the models for the training (In DB) and test sets (Out DB)

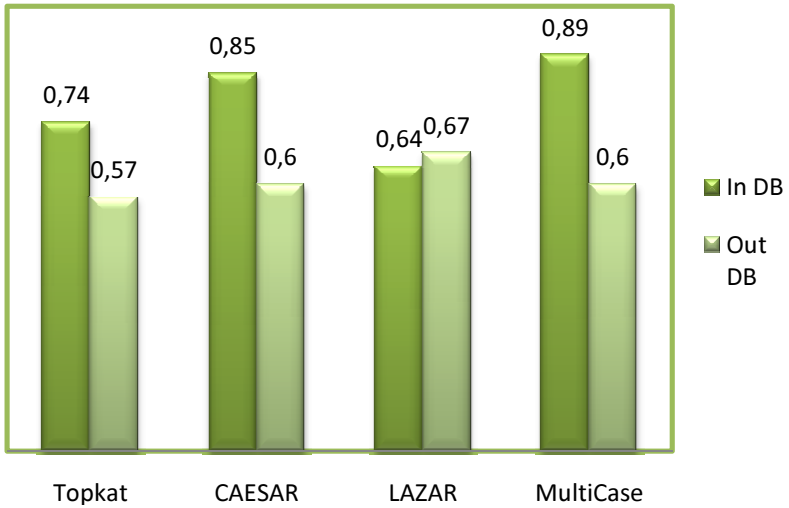
Sensitivity



Specificity



Accuracy



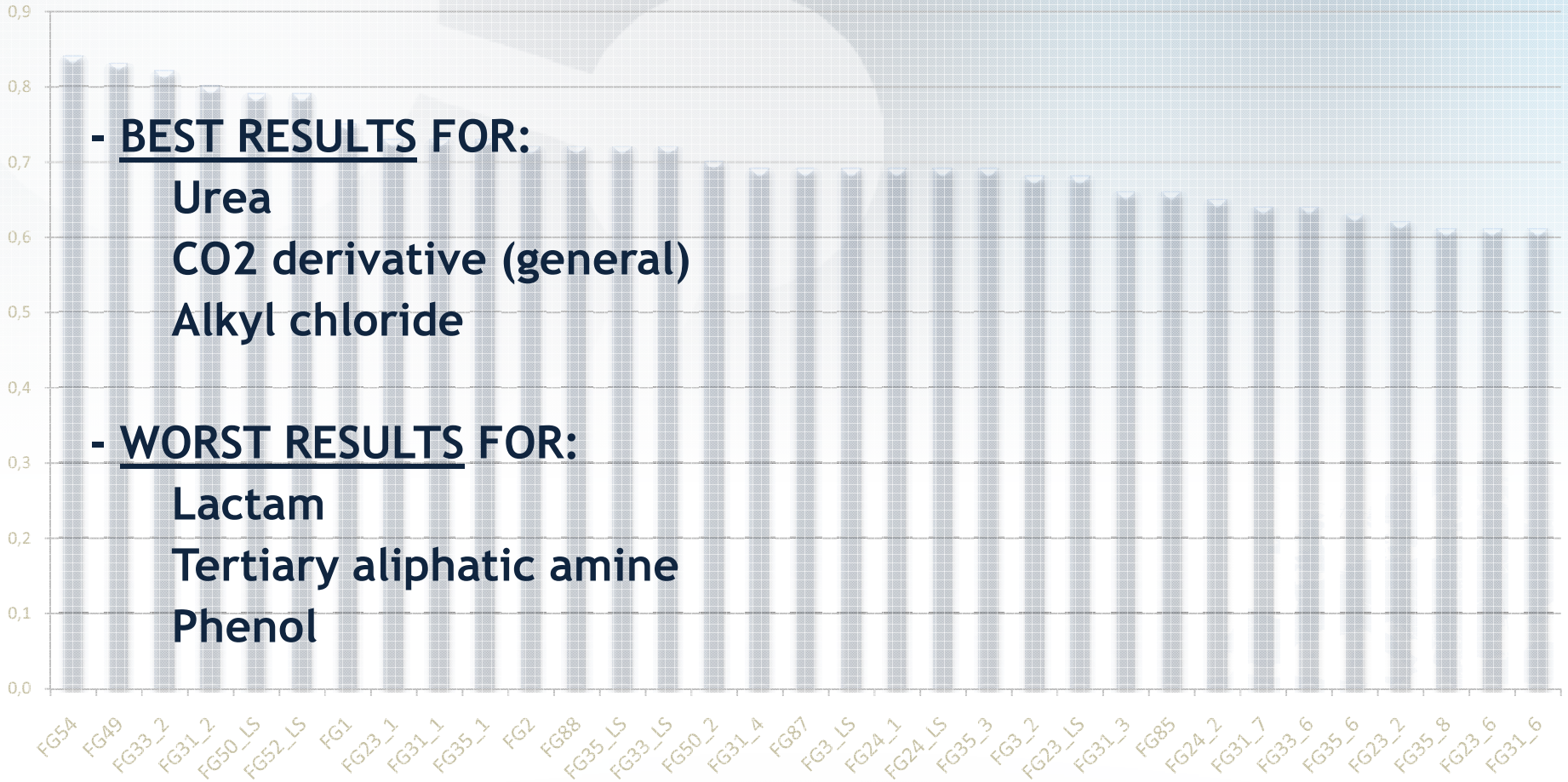
Carcinogenicity Results

Percentage of matched predictions stratified by mechanism of carcinogenicity

Mechanisms of carcinogenicity	TOXTREE	HAZARDEXPERT	DEREK	LAZAR	CAESAR	TOPKAT
Acylating direct acting	**77,8**	55,6	66,7	44,4	**77,8**	44,4
Alkylating direct acting	58,6	53,7	58,6	61,5	**70,9**	57,8
Alkylating indirect acting	79,2	70,7	78,8	75,7	**83,0**	67,6
Intercalating and DNA adduct forming Indirect acting	68,0	68,9	**76,7**	45,6	70,9	54,4
Aminoaryl DNA adducts forming Indirect acting	64,8	60,6	65,4	63,5	**74,3**	58,7
Non genotoxic	41,6	56,2	65,2	64,0	**71,9**	59,6
No Alerts	65,2	63,4	63,1	64,9	**70,5**	49,5

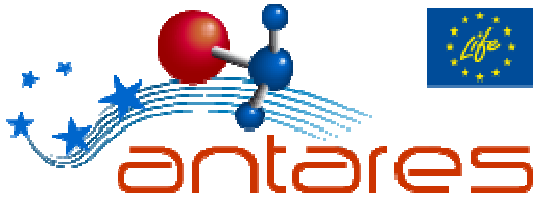
CARCINOGENICITY: Performance

Results per Chemical Classes (CAESAR)



- **BEST RESULTS FOR:**
  - Urea
  - CO2 derivative (general)
  - Alkyl chloride
- **WORST RESULTS FOR:**
  - Lactam
  - Tertiary aliphatic amine
  - Phenol

CALEIDOS starts where ANTARES ends



L FE08
ENV/IT/00435

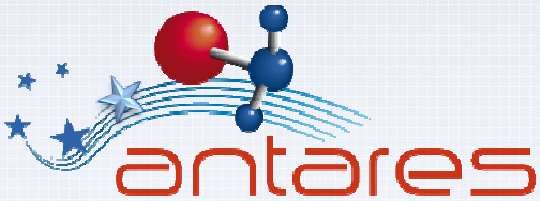
http://www.antares-life.eu/

IT ADDRESSED THE OVERALL PERFORMANCE OF QSAR METHODS AND IDENTIFIED RELIABLE QSAR MODELS USING GOOD QUALITY DATASETS

CALEIDOS WILL ADDRESS THE REGISTERED DATA

From ANTARES to VEGA

- Identification of the BEST MODELS
- Characterisation of the AD
- Integration of DIFFERENT MODELS
- Implementation into a UNIQUE PLATFORM
- Integration with READ ACROSS



VEGA

VEGA


VEGA
USE QSAR/ READ ACROSS
DOWNLOAD SOFTWARE
QSAR REGULATION & RESEARCH
ABOUT QSAR/ READ ACROSS
CONTRIBUTORS
ORCHESTRA
antares
CAESAR
Ministero della Salute
ISPRA
Our Vision
Our Mission
Our System
News & Updates
September 21
ANTARES list of predicting software for several REACH endpoints available
September 12
VEGA announced at the EUROTOX conference, Paris 2011
On site
November 23
Section HOW TO INTERPRET RESULTS added
November 23
VEGA website updated
Our Community
MARIO NEGRI
POLITECNICO DI MILANO
fera
The Food and Environment Research Agency
KnowledgeMiner Software
IFG
ARCHE
AARHUS UNIVERSITY
kode
EQUITOX

VEGA and the APPLICABILITY DOMAIN

The *different checks* done by VEGA for the definition of the Applicability Domain Index

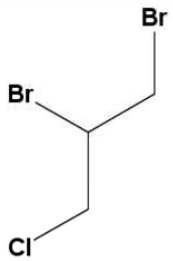
- **Visualisation of similar substances**
- **Similarity index** *(chemical; sub-indices)*
- **Chemiometric check** *(descriptor space)*
- **Atom centred-fragment** *(chemical)*
- **Check of the descriptor sensitivity** *(algorithm)*
- **Uncertainty** *(algorithm)*
- **Fragments for outliers** *(output space)*
- **Prediction Accuracy** *(output space)*
- **Prediction Concordance** *(tox exploration)*

APPLICABILITY DOMAIN INDEX

How the ADI information is visualized

1. Prediction Summary

Prediction for compound 1 (Molecule 1)



Prediction: Reliability:

Model assessment: Prediction is logBCF = 1.65, the result appears reliable. Anyhow, you should check it through the evaluation of the information given in the following sections.

Compound: 1
Compound SMILES: C(C(CBr)Br)Cl
Experimental value: -
Prediction: 1.65 [log(L/kg)]
Prediction: 45 [L/kg]
Prediction of model 1 (HM): 1.75 [log(L/kg)]
Prediction of model 2 (GA): 1.61 [log(L/kg)]
Structural Alerts: -
Calculated LogP: 2.96 [log units]
Reliability: Compound is in model Applicability Domain
Remarks for the prediction:
none

The *Applicability Domain Index* is summarized in one value, in top of the table of the Prediction Summary

All the *measured components contributing to the AD global index* are shown for an easy visualization of some potentially critical aspects.

3.2 Applicability Domain: Measured Applicability Domain Scores

Global AD Index
AD Index = 1
Explanation: predicted substance is into the Applicability Domain of the model.

Similar molecules with known experimental value
Similarity index = 0.981
Explanation: strongly similar compounds with known experimental value in the training set have been found.

Accuracy (average error) of prediction for similar molecules
Accuracy index = 0.19
Explanation: accuracy of prediction for similar molecules found in the training set is good.

Concordance with similar molecules (average difference between target compound prediction and experimental values of similar molecules)
Concordance index = 0.384
Explanation: similar molecules found in the training set have experimental values that agree with the target compound predicted value.

Maximum error of prediction among similar molecules
Max error index = 0.2
Explanation: the maximum error in prediction of similar molecules found in the training set has a low value, considering the experimental variability.

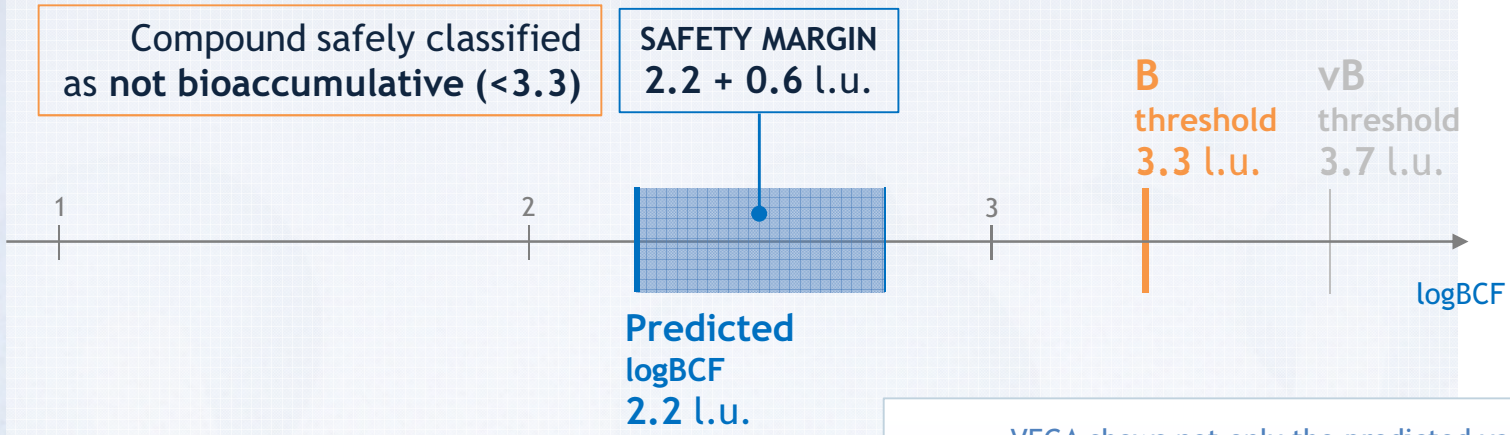
Atom Centered Fragments similarity check
ACF matching index = 1
Explanation: all atom centered fragment of the compound have been found in the compounds of the training set.

Descriptors noise sensitivity analysis
Noise Sensitivity = 0.913
Explanation: predictions has a good response to noise scrambling, thus shows a good reliability.

Model descriptors range check
Descriptors range check = true
Explanation: descriptors for this compound have values inside the descriptor range of the compounds of the training set.

ADEQUACY OF A MODEL

This chart shows (for BCF case) the predicted value together with its conservative confidence interval for safe classification



No. Comp. = 492		Exp. logBCF	
		nB	B/vB
Pred. logBCF	nB	359	0
	B/vB	60	73

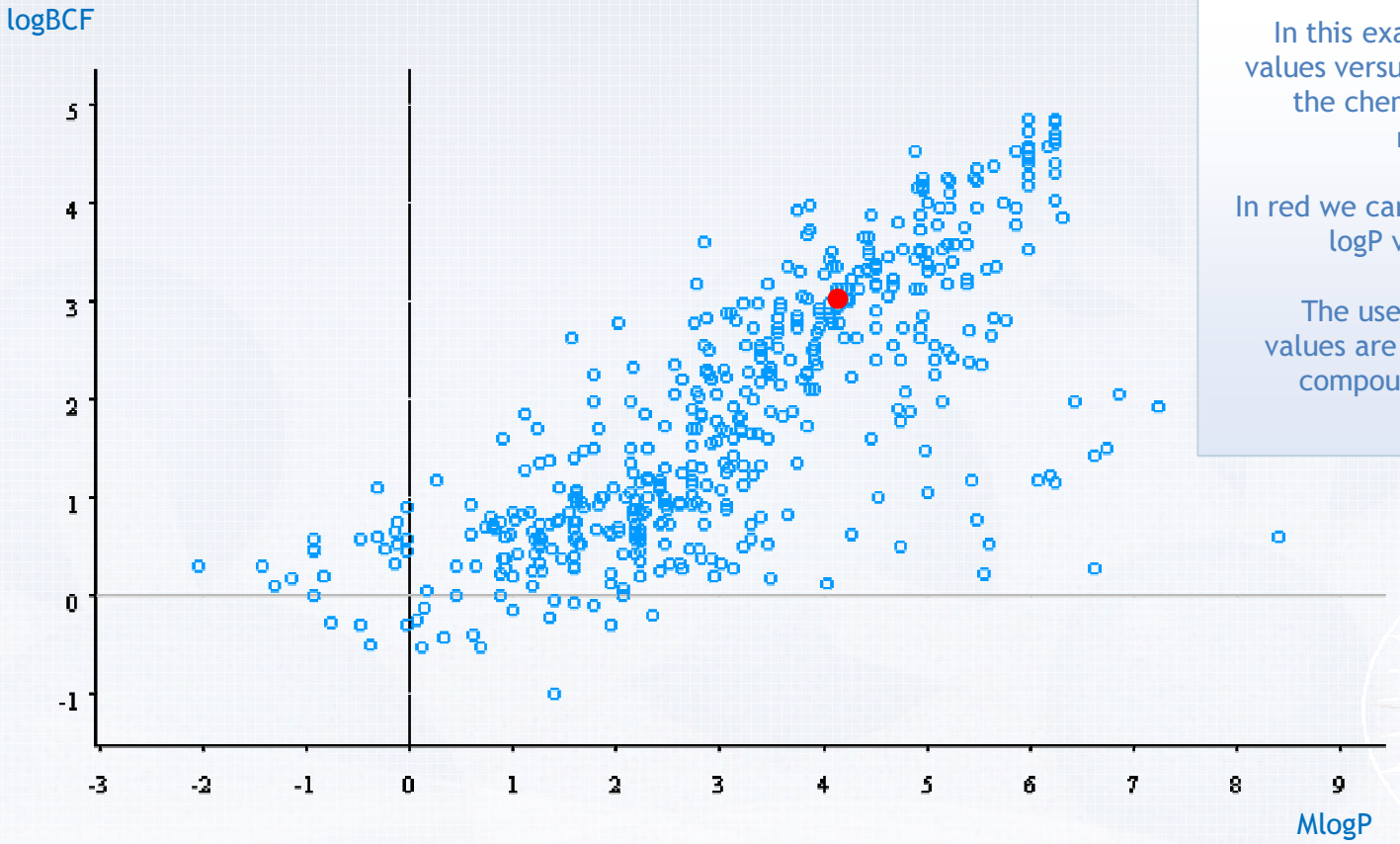
VEGA shows not only the predicted value (2.2 l.u.) but also its *uncertainty*, and how far it is from the threshold (3.3 l.u. for logBCF).

The *safety margin* (2.2 l.u. plus a conservative interval of 0.6 l.u.) is calculated specifically for each chemical, considering the ADI of the specific compound. In addition, it is determined in a way to provide no false negative prediction.

The confusion matrix verified on a set of 492 compounds is shown.

SUPPORTING DOCUMENTATION

VEGA provides *additional material* to support the prediction: DETAIL ON MOLECULAR DESCRIPTORS



In this example the experimental logBCF values versus the predicted logP values for the chemicals in the training set of the model are shown, as blue dots.

In red we can see the predicted logBCF and logP values of the target compound.

The user can evaluate if the predicted values are within the typical trend of the compounds, or if an unusual behaviour appears.

ANTARES Contribution to ANNEX XI

According to REACH Regulation (Annex XI)

a QSAR Model is VALID IF

1. the model is recognized *scientifically valid*;	• ANTARES contributed to assess model's validity
2. the substance is included in the *applicability domain* of the model;	• ANTARES provided results per chemical classes and MoA • VEGA improved ADI
3. results are adequate for *classification and labelling* and for *risk assessment*;	• VEGA introduced safety margin • Evaluation done in regression and classification
4. adequate *documentation of the methods* is provided.	• VEGA provided material (figures, framments, guidance to expert)

Other web sites and initiatives

http://www.orchestra-qsar.eu/

- Course
- E-book
- Movies
- Lessons
- Interviews

http://www.smart-reach.net/

Promoted by Italian authorities

THANK YOU
very much
for your kind
attention!