Emilio BENFENATI

Istituto di Ricerche Farmacologiche Mario Negri

QSAR and computational tools

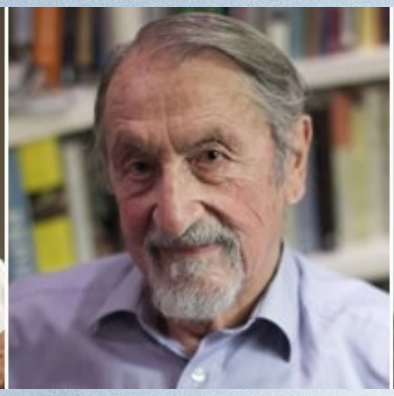
caleidos

Life

PROSIL

The Nobel Prizes in Chemistry 1998, 2013

The Nobel Prize in Chemistry 2013 has gone to Michael Levitt, Martin Karplus and Arieh Warshel, who “took the chemical experiments into cyberspace”


The second prize after the Nobel Prize in 1998 to John Pople and Alter Kohn for computational chemistry

Chemistry and cyberspace

All science is computer science *(New York Times)*

Millions of data to be processed, more and more common

In silico methods like the glue to integrate multiple evidences

Seven Reasons

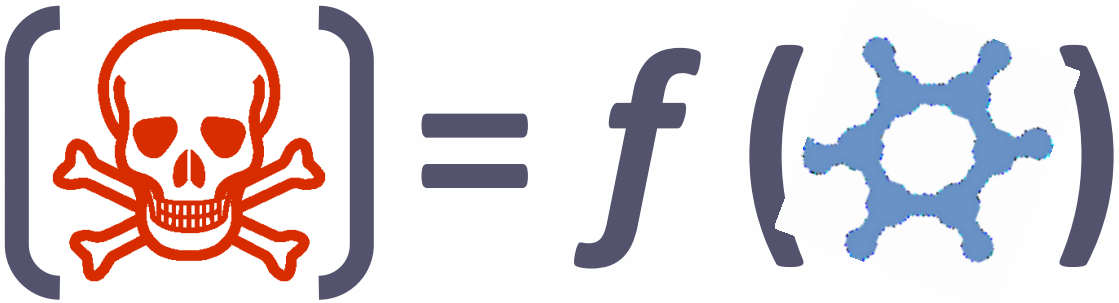
to use QSAR

1. Innovation *(also in view of millions of new data - ToxCast)*
2. Time for experiments
3. Occurrence of enough laboratories/resources
4. Reduction of costs
5. Use of animals
6. Prioritization needs
7. Pro-active approach for greener chemicals

(Q)SAR

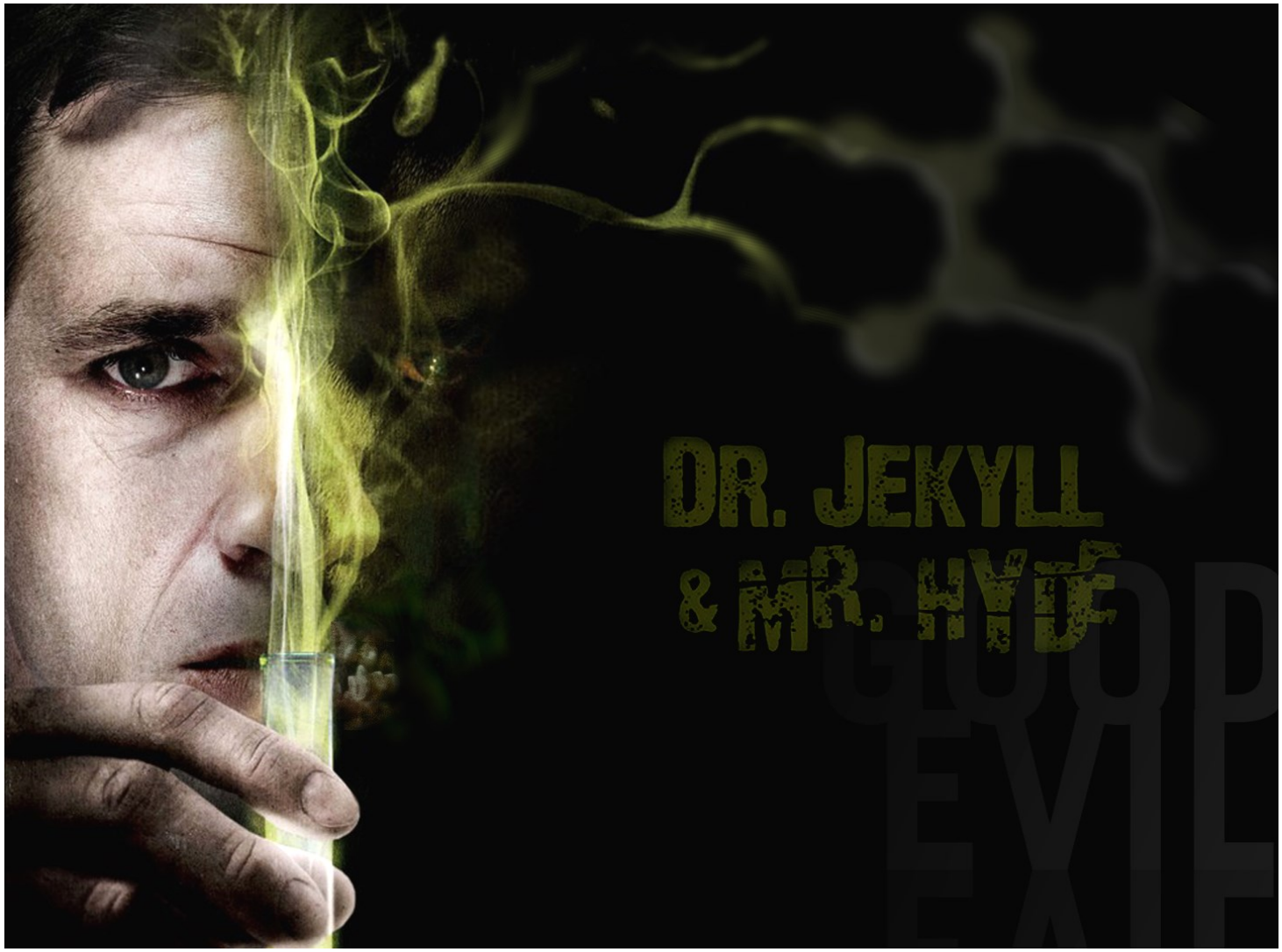
=

(Quantitative) Structure-activity relationship

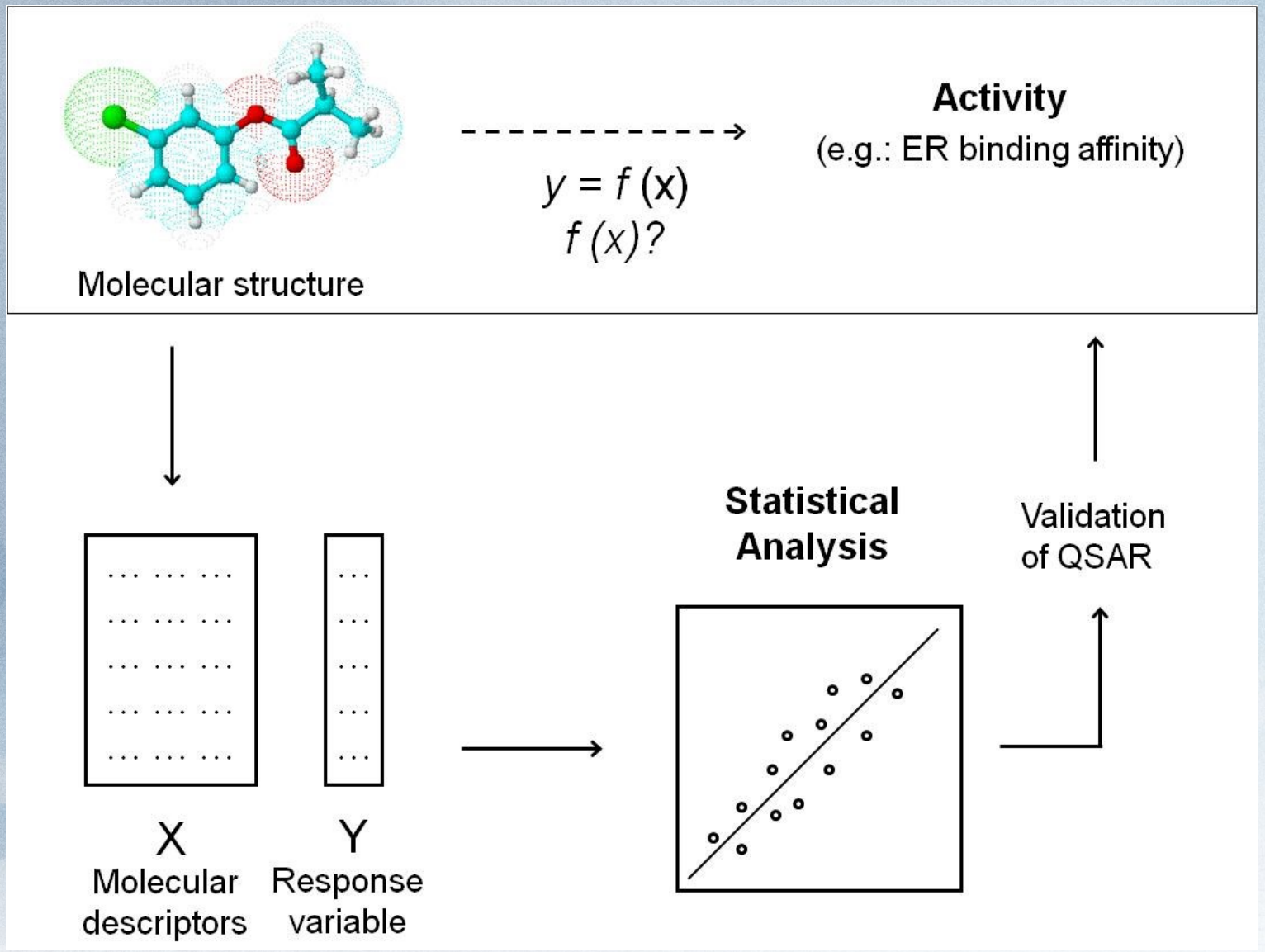


IN SILICO

CHEMICALS: GOOD and EVIL

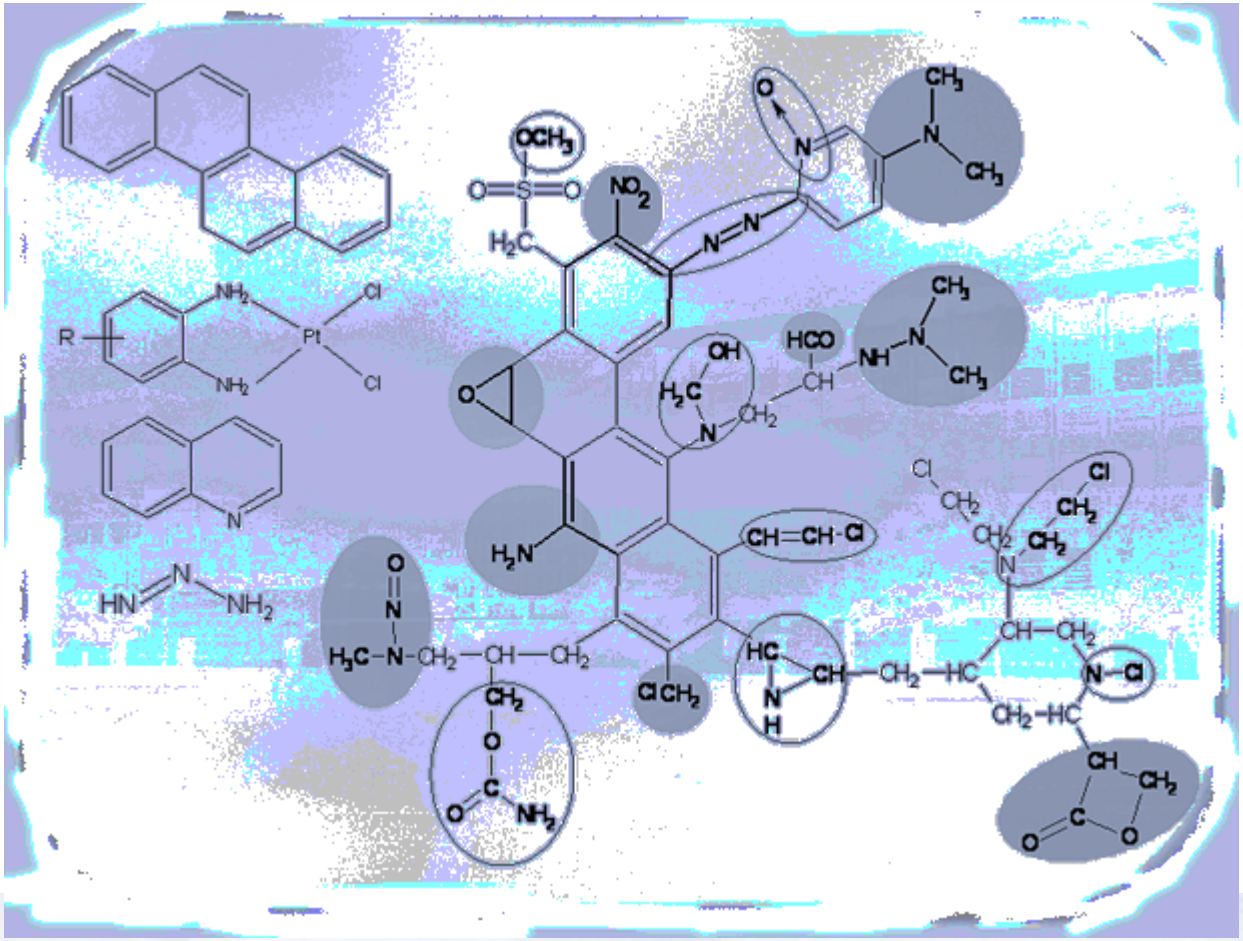


QSAR flow-chart

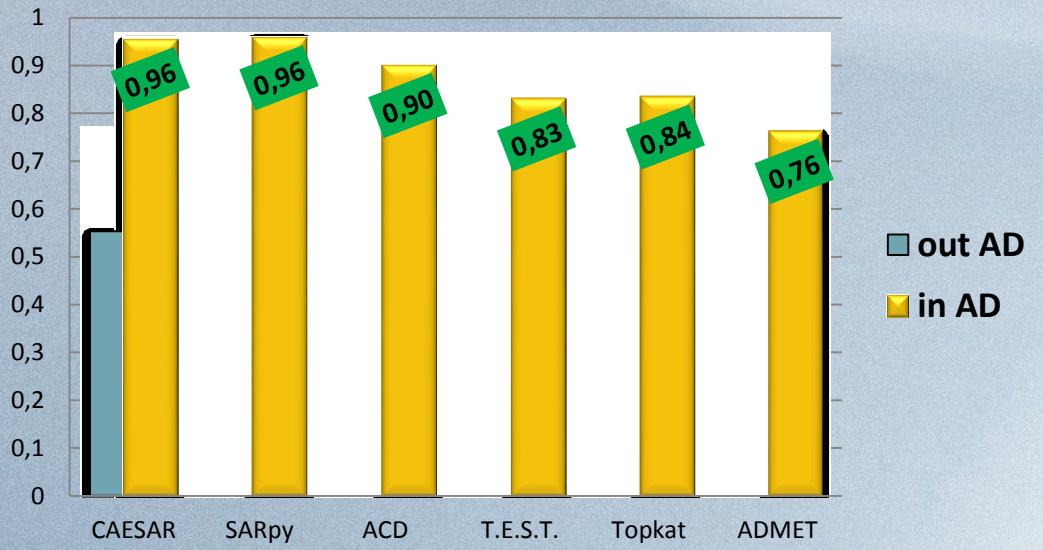


MOLECULAR FRAGMENTS AND ALERTS

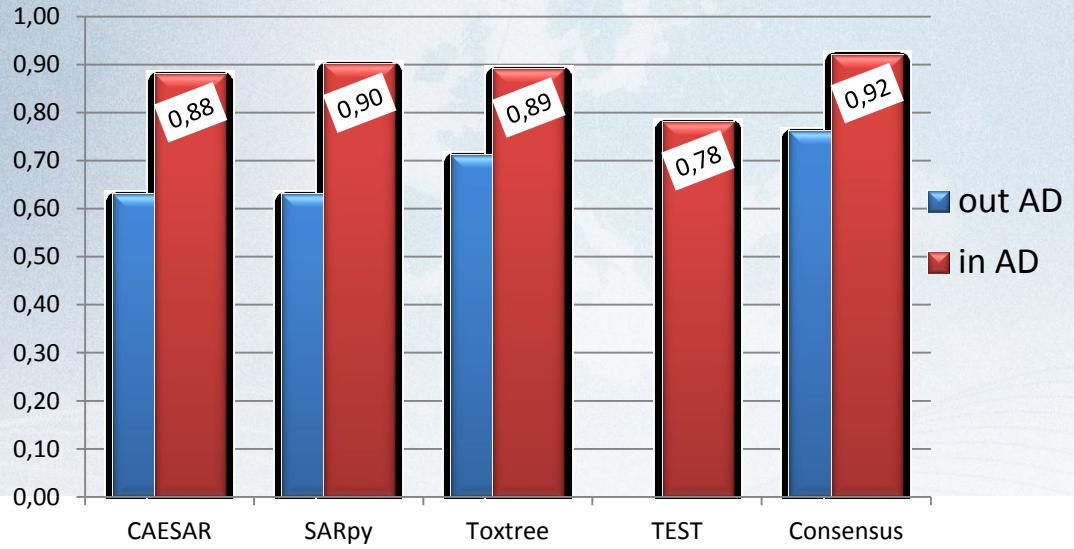
ASHBY identified RESIDUES for GENOTOXIC EFFECTS



MUTAGENICITY: Performance of QSAR models



$$Accuracy = \frac{TP + TN}{TP + FN + TN + FP}$$



Issues: Work in progress

- *Max accuracy for carcinogenicity models: 0.75 (Toxtree, in VEGA)*
- *Max accuracy for devtox models: 0.78 (SARpy + P&G, in VEGA), but MCC 0.24 (false negatives)*
- Problem 1: Complexity of the endpoints
- Problem 2: Lack of data

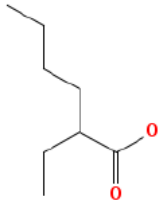

VEGA
USE QSAR/ READ ACROSS
DOWNLOAD SOFTWARE
QSAR REGULATION & RESEARCH
ABOUT QSAR/ READ ACROSS
CONTRIBUTORS
ORCHESTRA
antares
CAESAR
Ministero della Salute
MINISTERO DELL'AMBIENTE
ISPRA
Istituto Superiore per la Protezione e la Ricerca Ambientale
Our Vision
Our Mission
News & Updates
September 21
ANTARES list of predicting software fo
several REACH endpoints available
September 12
VEGA announced at the EUROTOX
conference, Paris 2011
MARIO NEGRI
ISTITUTO DI RICERCHE FARMACOLOGICHE
POLITECNICO DI MILANO
VegaNIC - VEGA Non-Interactive Client - version 1.1.0
INSERT
SELECT
EXPORT
PREDICT
Select models
VEGA NIC
Tox
Ecotox
Environ
Phys-Chem
Mutagenicity (Ames test) model (CAESAR) - v. 2.1.13
Mutagenicity (Ames test) model (SarPy/IRFMN) - v. 1.0.7
Mutagenicity (Ames test) model (ISS) - v. 1.0.2
Mutagenicity (Ames test) model (KNN/Read-Across) - v. 1.0.0
Carcinogenicity model (CAESAR) - v. 2.1.9
Carcinogenicity model (ISS) - v. 1.0.2
Developmental Toxicity model (CAESAR) - v. 2.1.7
Developmental/Reproductive Toxicity library (PG) - v. 1.0.0
Relative Binding Affinity model (IRFMN) - v. 1.0.1
Skin Sensitisation model (CAESAR) - v. 2.1.6
All models

www.vega-qsar.eu

VEGA

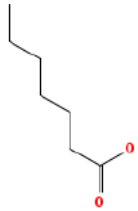
Mutagenicity (Ames test) m

3.1 Applicability Domain:
Similar Compounds, with Predicted and



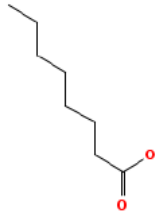
CAS: 149-57-5
Dataset id: 536 (Training
SMILES: O=C(O)C(CC)C
Similarity: 0.989

Experimental value: NON-
Predicted value: NON-Mu



CAS: 111-14-8
Dataset id: 86 (Training s
SMILES: O=C(O)CCCCC
Similarity: 0.946

Experimental value: NON-
Predicted value: NON-Mu



CAS: 124-07-2
Dataset id: 418 (Training
SMILES: O=C(O)CCCCC
Similarity: 0.941

Experimental value: NON-
Predicted value: NON-Mu

VEGA

Mutagenicity (Ames test) model (CAESAR) 2.1.13 pa

3.2 Applicability Domain:
Measured Applicability Domain Scores

Global AD Index
AD index = 0.978
Explanation: the predicted compound is into the Applicability Domain of the model.

Similar molecules with known experimental value
Similarity index = 0.956
Explanation: strongly similar compounds with known experimental value in the training set have been found.

Accuracy of prediction for similar molecules
Accuracy index = 1
Explanation: accuracy of prediction for similar molecules found in the training set is good.

Concordance for similar molecules
Concordance index = 1
Explanation: similar molecules found in the training set have experimental values that agree with the predicted value.

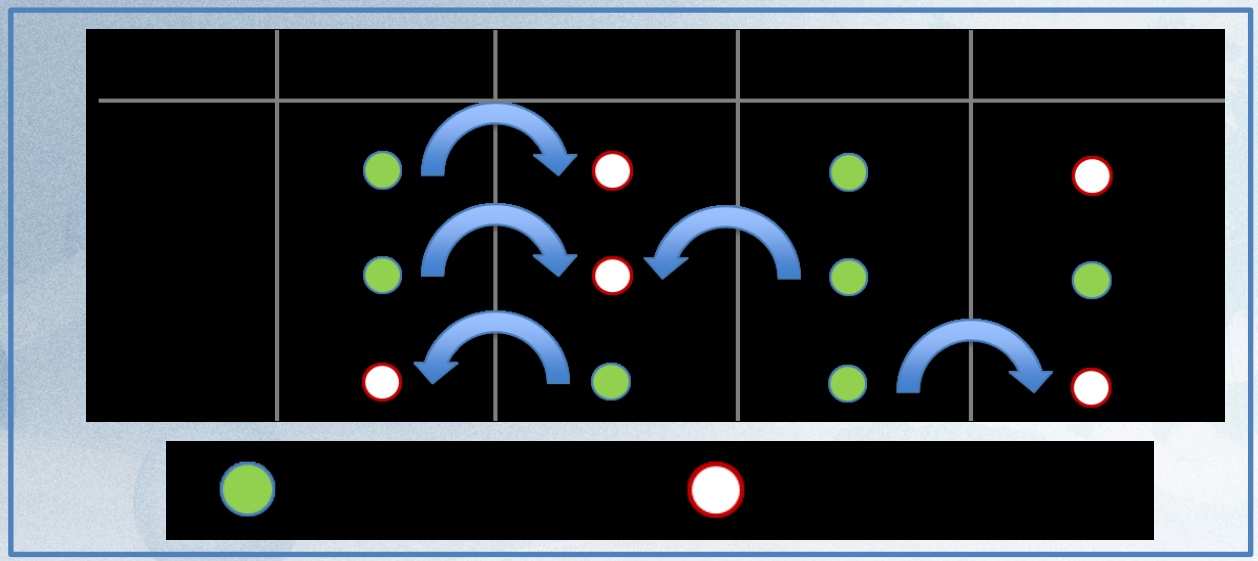
Model's descriptors range check
Descriptors range check = True
Explanation: descriptors for this compound have values inside the descriptor range of the compounds of the training set.

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

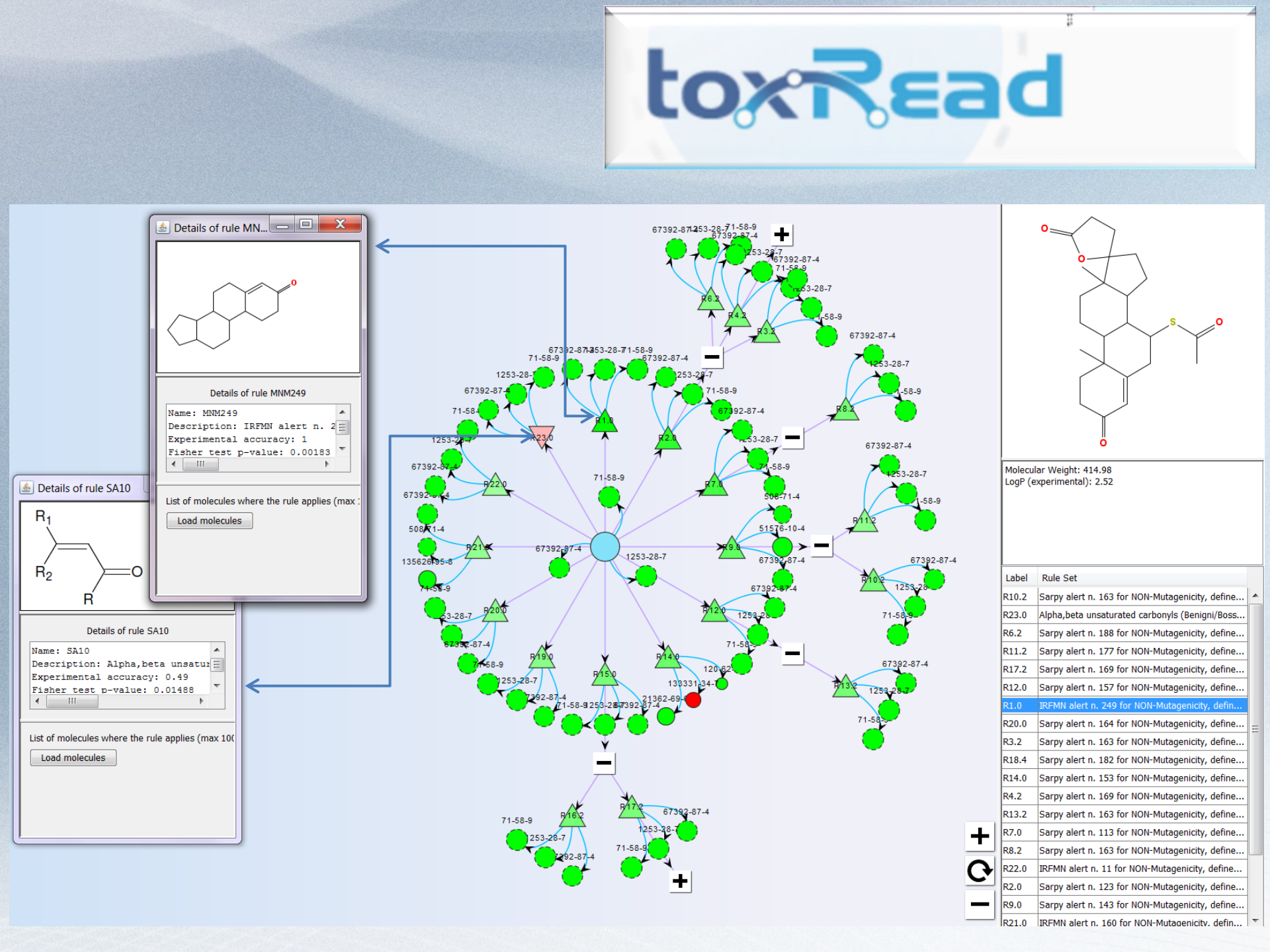
Read-across

Read-across: a correlation or relationship between two separate things

From a chemical point of view: Read-across is a method for data-gap filling where information from one or more chemicals is used to predict the same endpoint for a target chemical



www.toxgate.eu


toxRead
Details of rule MN...
Details of rule MNM249
Name: MNM249
Description: IRFMN alert n. 2
Experimental accuracy: 1
Fisher test p-value: 0.00183
List of molecules where the rule applies (max
Load molecules
Details of rule SA10
Details of rule SA10
Name: SA10
Description: Alpha,beta unsatu
Experimental accuracy: 0.49
Fisher test p-value: 0.01488
List of molecules where the rule applies (max 10
Load molecules
Molecular Weight: 414.98
LogP (experimental): 2.52
Label
Rule Set
R10.2 Sarpy alert n. 163 for NON-Mutagenicity, define...
R23.0 Alpha,beta unsaturated carbonyls (Benigni/Boss...
R6.2 Sarpy alert n. 188 for NON-Mutagenicity, define...
R11.2 Sarpy alert n. 177 for NON-Mutagenicity, define...
R17.2 Sarpy alert n. 169 for NON-Mutagenicity, define...
R12.0 Sarpy alert n. 157 for NON-Mutagenicity, define...
R1.0 IRFMN alert n. 249 for NON-Mutagenicity, defin...
R20.0 Sarpy alert n. 164 for NON-Mutagenicity, define...
R3.2 Sarpy alert n. 163 for NON-Mutagenicity, define...
R18.4 Sarpy alert n. 182 for NON-Mutagenicity, define...
R14.0 Sarpy alert n. 153 for NON-Mutagenicity, define...
R4.2 Sarpy alert n. 169 for NON-Mutagenicity, define...
R13.2 Sarpy alert n. 163 for NON-Mutagenicity, define...
R7.0 Sarpy alert n. 113 for NON-Mutagenicity, define...
R8.2 Sarpy alert n. 163 for NON-Mutagenicity, define...
R22.0 IRFMN alert n. 11 for NON-Mutagenicity, define...
R2.0 Sarpy alert n. 123 for NON-Mutagenicity, define...
R9.0 Sarpy alert n. 143 for NON-Mutagenicity, define...
R21.0 IRFMN alert n. 160 for NON-Mutagenicity, defin...

CONCLUSIONS

- *Computational models as support to human experts*
- *Navigation within data and reasoning*
- *No conflict between "computer" and man*
- *Multiple in silico approaches*
- *Integrating multiple approaches (weight of evidence)*
- *Comparison with the experimental uncertainty/vairability*