An AI-powered electronic lab notebook for sustainable chemistry

University of Nottingham
UK | CHINA | MALAYSIA

Royal Academy of Engineering

Joseph Heeley, Ton Blackshaw, Joseph C. Davies, Zakariyyaa Siddiq, Madeleine Parker, Jonathan D. Hirst

An Unsustainable Data Problem

Improving sustainability is an important consideration in the chemical industry that has both environmental and societal impacts.

Though guidelines such as the 12 Principles of Green Chemistry[1] have been developed, sustainable chemistry remains a formidable challenge.

The application of artificial intelligence could provide a powerful solution to identify more sustainable choices in the lab.

However, many chemists still record experimental data in paper notebooks, which is not sustainable from an environmental or data-driven perspective.

Our Solution: AI$_4$Green ELN

AI4Green[2] is a free-to-use electronic laboratory notebook (ELN) that encourages users to make sustainable choices.

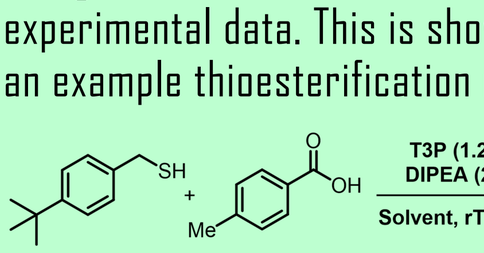
Key Features:

- Automatic molar calculations
- Free data export
- Automatic hazard and sustainability assessment
- Cloud-based data backup
- Professional COSHH matrix
- Seamless and secure collaboration

Scan here to try AI$_4$Green

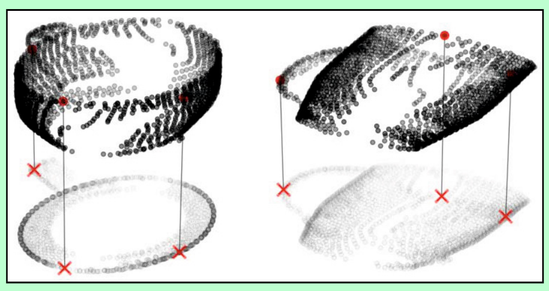
Selecting Sustainable Solvents

Interactive principal component analysis



Principal Component Analysis (PCA) is a statistical technique that reduces the dimensionality of data while still retaining its variance.

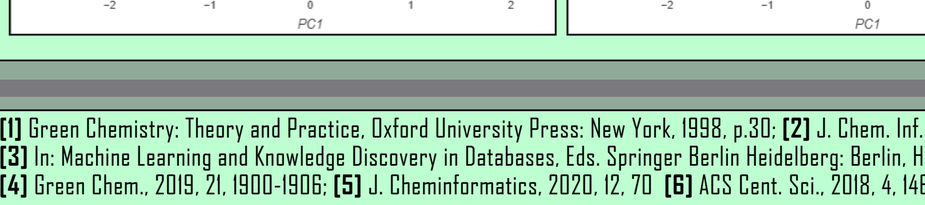
Interactive PCA[3] allows users to define positions of data points by dragging them together and updating the positions of the remaining points.



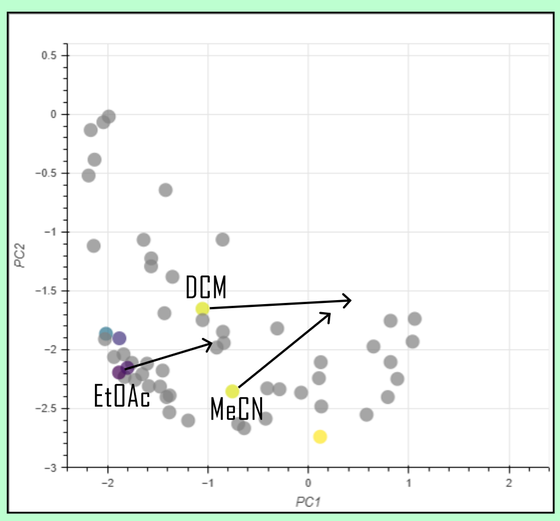
The Solvent Surfer

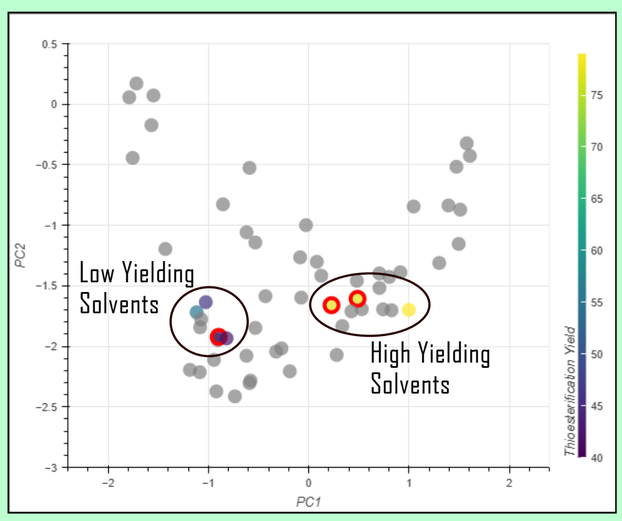
AI$_4$Green's Solvent Surfer tool exploits interactive kernel PCA of 16 physicochemical descriptors.

Using the Solvent Surfer, solvent maps can be tailored to closely match experimental data. This is shown in the case study of a solvent screen for an example thioesterification reaction.[4]



Solvent	DMC	EtOAc	CPME	2-MeTHF	MeCN	DCM	DMF
CHEM21							
% Yield	55	40	46	42	77	77	79

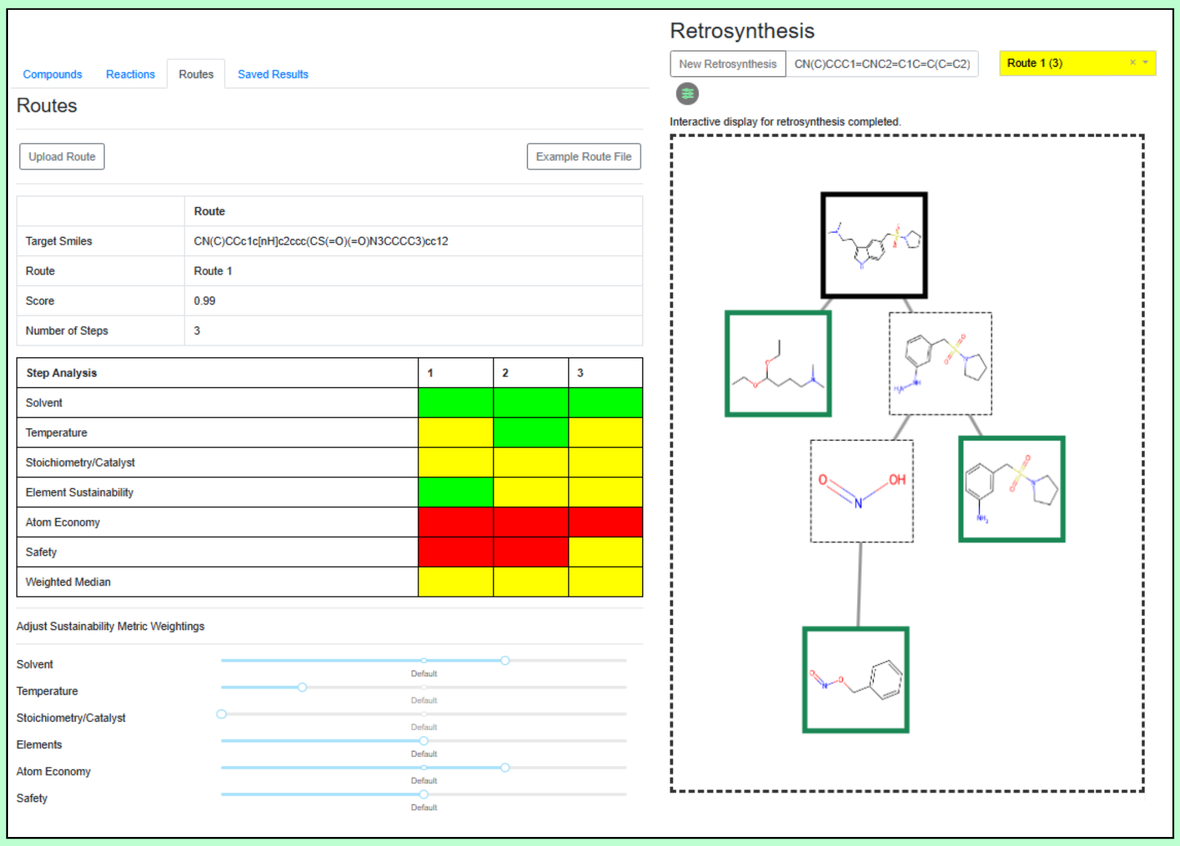




Sustainable Retrosynthesis

The retrosynthesis tool in AI$_4$Green combines AIZynthFinder[5] and ASKCOS[6] for prediction of retrosynthesis and reaction conditions.

All routes are evaluated according to important sustainability metrics. Users can adjust weights of important features to find the most suitable route.



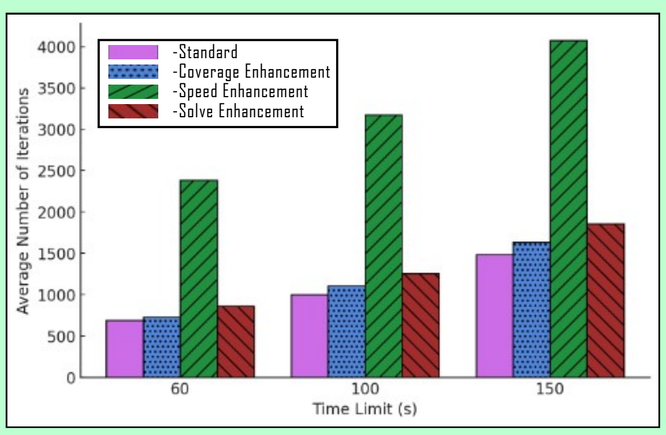
Enhancing Retrosynthesis

AIZynthFinder uses a Monte Carlo Tree Search (MCTS) to efficiently search through 12 million reaction templates extracted from the USPTO dataset.

The implementation in AI$_4$Green exploits three novel enhancements to the MCTS, which alter the search based on the current state.

These enhancements give improvements to speed, search coverage and solve rate compared to the standard MCTS.

Enhancement	Solve %	Routes per Molecule	Seconds per Molecule
None	86.7	31.3	25.5
Coverage	86.7	32.5	24.6
Speed	81.6	25.6	11.9
Solve	86.9	32.1	18.6



[1] Green Chemistry: Theory and Practice, Oxford University Press: New York, 1998, p.30; **[2]** J. Chem. Inf. Model. 2023, 63, 2895–2901; **[3]** In: Machine Learning and Knowledge Discovery in Databases, Eds. Springer Berlin Heidelberg: Berlin, Heidelberg, 2014, pp 501–516; **[4]** Green Chem., 2019, 21, 1900-1906; **[5]** J. Cheminformatics, 2020, 12, 70 **[6]** ACS Cent. Sci., 2018, 4, 1465–1476