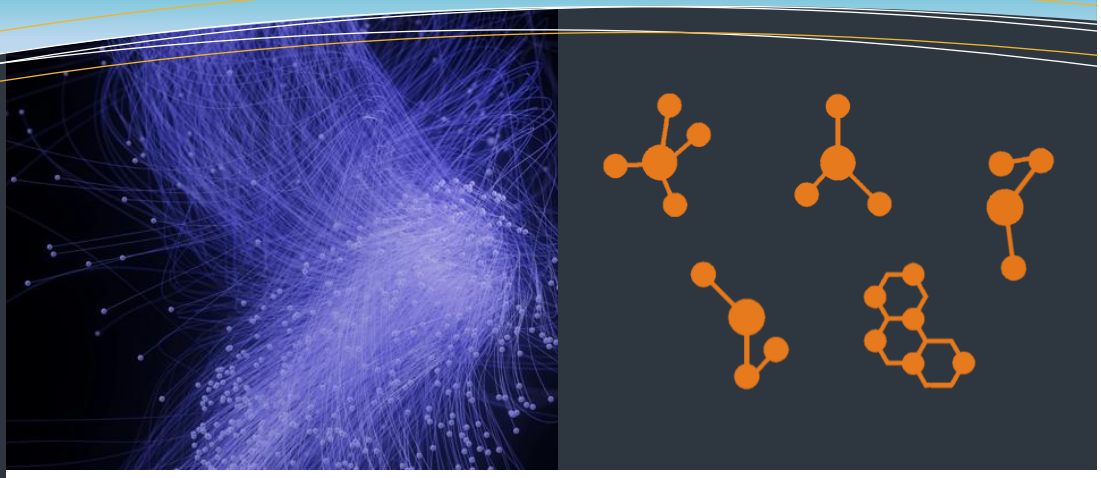
VOLUME
44
June
2026

GT
GERMAN SOCIETY OF
TOXICOLOGY

Poison of the month



Chemical Space was selected as Poison of the Month for June by the Working Group Computational Toxicology

…because the working group addresses the question of how “poison” can be located within chemical space. The focus is therefore less on a single toxic substance and more on the search for “toxic” patterns.

What is Chemical Space?

Chemical Space refers to all molecules that exist or could, in principle, be synthesized. It is not a physical space, but a theoretical, high-dimensional property space of chemical structures. Its size is almost unimaginable: for thermodynamically stable molecules, around 10^{440} possible structures have been estimated. Even for smaller, drug-like molecules, about 10^{60} structures are assumed.

By contrast, the fraction actually known, synthesized, or tested is vanishingly small. Databases therefore capture only selected parts of this space, for example “commercially available chemicals”.

Molecules are positioned according to properties such as size, polarity, lipophilicity, functional groups, ring systems, or molecular fingerprints. Similar molecules are located closer to each other than dissimilar ones. Because this space has many dimensions, mathematical methods project it onto two- or three-dimensional maps. These simplified maps show which substance classes are well studied, and where toxicological data remain sparse.

Chemical Space in Toxicology: Mapping Toxic Patterns

The concept of Chemical Space emerged from cheminformatics, drug discovery, and computational chemistry. It describes chemical structures as points within a property space, where proximity and distance reflect similarity or dissimilarity. Through new databases, molecular fingerprints, high-throughput screening, and machine learning, this space is continuously expanded and explored for different scientific and regulatory questions.

For computational toxicology, this concept opens new opportunities to systematically use toxicological data for safety assessment. By linking chemical structural information with toxicological endpoint data from studies, it becomes possible to investigate whether regions of chemical space can be described in which substances with certain toxicological properties occur more frequently, and whether these regions can be distinguished from areas that predominantly contain less concerning substances. Such “toxic subspaces” or “safe-zone subspaces” are not fixed geographical zones, but model-dependent regions derived from structural features, physicochemical properties, and toxicological endpoints.

The scientific reliability of such distinctions depends largely on the underlying data basis. QSAR models, read-across, structural alerts, and machine-learning approaches rely on existing data and chemical similarity. If a substance lies within a well-described region, predictions can be supported by comparable reference substances. If, however, it is located in a sparsely populated region, results may appear formally plausible even though their empirical basis is weak. Chemical Space analyses therefore reveal not only possible toxicological patterns, but also the limits of prediction.

In this way, Chemical Space can help to avoid animal testing in a more targeted manner without weakening the scientific robustness of safety

Deutsche Gesellschaft für Toxikologie

VOLUME
44
June
2026

GT
GERMAN SOCIETY OF
TOXICOLOGY

Poison of the month

Chemical Space in Regulatory Application

In a recent study, the German Federal Institute for Risk Assessment (BfR) investigated how well the chemical space of pesticide- and biocide-relevant substances is covered by public genotoxicity data. For this purpose, 18 public pesticide and biocide databases were merged, structurally standardized, and reduced to 4,826 clearly identifiable substances. These were compared with 19,897 substances from 19 public genotoxicity databases for which at least one data point from at least one genotoxicity test was available.

The overlap was assessed at several levels: physicochemical properties, functional groups, molecular scaffolds, and exact chemical identity. For the Ames test, the physicochemical properties of pesticides were comparatively well covered by existing genotoxicity data. For the in vivo chromosomal aberration test, however, large parts of the physicochemical pesticide space lacked corresponding data. At the level of exact chemical identity, the overlap was limited: only around 11% of the genotoxicity substances were also found in pesticide datasets; for approximately 2,700 pesticide- and biocide-relevant substances, no public genotoxicity data were available.

Particularly for pesticide regulation, where non-testing methods such as read-across and QSAR are increasingly being used, this analysis shows where existing data can support predictions, and where the data basis is not yet sufficient for regulatory conclusions.

assessment. In well-characterized regions, existing data can be used more effectively, substances can be grouped appropriately, and unnecessary repeat testing can be avoided. In sparsely populated regions, by contrast, the analysis indicates where new data would be particularly relevant. Testing needs can thus be justified in a more data-driven, endpoint-specific, and mechanistically informed way.

At the same time, the concept of a “toxic space” must not be equated with risk. Toxicity is endpoint-dependent and is influenced by dose, bioavailability, metabolism, and exposure. Chemical Space therefore initially describes hazard-relevant structural and property patterns. For a complete risk assessment, exposure and use scenarios must additionally be considered.

A recent study on the coverage of pesticide- and biocide-relevant substances by public genotoxicity data illustrates how Chemical Space analyses can be applied to toxicological questions. The background is that non-testing methods such as read-across and QSAR are increasingly used in pesticide regulation. However, their reliability depends on whether the substance domain under assessment is sufficiently represented by suitable reference substances and test data. For pesticide- and biocide-relevant substances, the study therefore provides important guidance as it shows where predictions of genotoxicity are supported by existing data and where additional data are needed to substantiate regulatory conclusions.

The potential of Chemical Space thus lies not only in improving predictions, but above all in making their limits visible. This is a key prerequisite for robust, transparent, and increasingly animal-free toxicological assessments.

.

By Ute Haßmann

Literature and links:

- Comparative Chemical Space Analysis of Pesticides and Substances with Genotoxicity Data | Chemical Research in Toxicology
- Exploring Chemical Space for Drug Discovery Using the Chemical Universe Database | ACS Chemical Neuroscience
- Exploring Chemical Space with Machine Learning
- Chemical Space - an overview | ScienceDirect Topics
- Foto by Max Petrunin on Unsplash