**////Title: Exploring the Vast History of Chemical Space**

**////Standfirst:**

Chemical space, a term referring to the catalogue of all known molecules, has been constantly expanding for over 200 years. Our understanding of this space is based on the work of countless researchers over the years. Dr Guillermo [Gee-yehr-moe] Restrepo and his team at the Max Planck Institute for Mathematics in the Sciences, Germany, have been analysing trends in the history of chemistry, to gain an understanding of how the field became what it is today.

**////Main text:**

In the world of chemistry, the discovery of new molecules is a milestone for many scientists. Chemical scientists dedicate significant time and effort to the pursuit of these new compounds – both those found in nature, and those synthesised in the laboratory. Every new molecule that is discovered is recorded, leading to a vast library of data.

This catalogue of molecules, known as ‘chemical space’, grows every year. Today, in the information age, we have a perfect opportunity to look back at how far we have come, and evaluate progress in the field.

Dr Guillermo Restrepo and his team at the Max Planck Institute for Mathematics in the Sciences have been exploring the history of chemistry in order to better understand how the field has developed over time, and analyse how new chemical discoveries have come into being. Through this, it is hoped that we will understand how a field changes over time, which may also prepare us for the future.

…

The team’s work revolved around use of the Reaxys [ree-axis] database. Reaxys draws information from over 16400 scientific journals and spans over 200 years. This data covers over 21 million compounds and 43 million reactions.

From analysing the data in Reaxys, Dr Restrepo found that every year, the number of known chemical compounds has increased in an exponential fashion by 4.4%, and has done each year since 1800. Then, he was able to break down the growth of chemical space into three distinct historical regimes.

The first of these, referred to as the proto-organic regime, occurred up until 1860. During this time, discoveries were occurring less consistently year to year. Although many compounds involved metals, the important role of organic compounds started to become evident. Most newly discovered compounds were derived from animals and plants, although chemical synthesis still made up a significant proportion of the new compounds discovered.

…

The next regime is where organic chemistry really dominated. New organic compounds were being discovered more regularly year-to-year from 1861 until 1980. Dr Restrepo explains that this was partially due to the introduction and a better understanding of the theories that form the foundation of organic chemistry, which in turn led to a more calculated and successful approach to chemical synthesis.

…

The third and final era that the team identified is called the organometallic regime. Since 1981, there has been a noticeable drop in the number of purely organic compounds being discovered, in favour of those that consist of varied elements.

The use of nitrogen in synthesis is now higher than ever, and metal elements have finally found more relevance. Compounds that include silicon, despite being very rare previously, have become increasingly common during this period. While the number of new organic molecules discovered every year has dropped a little, there is now far more variety in terms of the types of compounds than ever before.

Nevertheless, the tradition of organic chemistry research remains a major force in chemistry. Since 1890, most reported compounds are made up of combinations of carbon, hydrogen, nitrogen and oxygen.

…

Of course, science never occurs in a vacuum, so to speak. It is important to understand how larger scale events in the outside world impact how scientific research is conducted. Dr Restrepo’s team investigated this by analysing the patterns in the growth of chemical space, before, during and after the two world wars of the 20th century.

As one might expect, there was a very significant drop in the number of new compounds being discovered during both world wars. In fact, Dr Restrepo suggests that, based on the lowest reporting trends in the data, the first world war set chemistry back by a massive 37 years, while the second regressed it by a substantial 16 years.

However, there was a silver lining. Directly after both wars, there was a significant spike in the number of new molecules being discovered, as if to make up for lost time. In fact, the years following the first world war showed the highest rate of growth for chemistry in all of recorded history. This meant that overall, despite the challenges that the wider world had to endure, the world of chemistry managed to bounce back.

…

There were also significant events in the history of science that influenced the direction of chemical research. Previously, the synthesis of urea in 1828 was considered to be the first real breakthrough in terms of chemical synthesis, which paved the way for all artificially developed compounds that came after it. While this was undoubtedly a very significant milestone in the field, Dr Restrepo wanted to investigate whether this was truly the beginning of chemical synthesis as we know it.

By looking into the data, he found that at the start of the 19th century, just under half of all compounds discovered each year were found through synthesis. This means that even before the fateful synthesis of urea, scientists were already synthesising a large proportion of newly discovered molecules. This continued to increase year by year, until the start of the 20th century, where around 90% of species were discovered through chemical synthesis. This fact has remained the same since then, showing that such a huge amount of work goes into developing compounds that can be created in the laboratory.

…

Dr Restrepo has thoroughly investigated the data for over 200 years of chemical research, summarised interesting trends in the history of chemical space, and investigated how these came about, due to both events in the field and the wider world. This has given us a fascinating insight into how our understanding of chemistry has developed over time, and how it may continue to change in the future.

This SciPod is a summary of the paper ‘Exploration of the chemical space and its three historical regimes’, from PNAS. <https://doi.org/10.1073/pnas.1816039116>