

In a world where annual economic losses from corrosion surpass 2.5 trillion US Dollars, the quest for <u>corrosion-resistant</u> alloys and protective coatings is unbroken. Artificial intelligence (AI) is playing an increasingly pivotal role in designing new alloys. Yet, the predictive power of AI models in foreseeing corrosion behaviour and suggesting optimal alloy formulas has remained elusive.

Scientists of the Max-Planck-Institut für Eisenforschung (MPIE) have now developed a machine learning model that enhances the predictive accuracy by up to 15% compared to existing frameworks. This model uncovers new, but realistic <u>corrosion-resistant</u> alloy compositions. Its distinct power arises from fusing both numerical and textual data. Initially developed for the critical realm of resisting pitting corrosion in <u>high-strength</u> alloys, this model's versatility can be extended to all alloy properties. The researchers published their latest results in the journal *Science Advances*.

Merging texts and numbers

"Every alloy has unique properties concerning its corrosion resistance. These properties do not only depend on the alloy composition itself, but also on the alloy's manufacturing process. Current machine learning models are only able to benefit from numerical data. However, processing methodologies and experimental testing protocols, which are mostly documented by textual descriptors, are crucial to explain corrosion" explains Kasturi Narasimha Sasidhar, lead author of the publication and former postdoctoral researcher at MPIE. The researcher team used language processing methods, akin to ChatGPT, in combination with machine learning (ML) techniques for numerical data and developed a fully-automated natural language processing framework. Moreover, involving textual data into the ML framework allows to identify enhanced alloy compositions resistant to pitting corrosion. "We trained the deep-learning model with intrinsic data that contain information about corrosion properties and composition. Now, the model is capable of identifying alloy compositions that are critical for corrosion-resistant materials even if the individual elements were not fed initially into the model," says Michael Rohwerder, co-author of the publication and head of the group Corrosion at MPIE.

Pushing boundaries: automated data mining and image processing

In the recently devised framework, Sasidhar and his team harnessed manually gathered data as textual descriptors. Presently, their objective lies in automating the process of data mining and seamlessly integrating it into the existing framework. The incorporation of microscopy images marks another milestone, envisioning the next generation of AI frameworks that converge textual, numerical, and image-based data.