

Projekt „Hand-drawn Chemical Structures“

Autoren: Saaruky Chanthirakanthan, Michelle Cassano, Zeynep Dagtekin, Safiye-Sarah Kantar, Fatma Zehra Sevindik

Gemeinsam mit der Arbeitsgruppe aus der Friedrich-Schiller-Universität Jena von Prof. Dr. Christoph Steinbeck und der Westfälischen Hochschule, geleitet von Prof. Dr. Achim Zielesny wurde ein Lernsystem entwickelt. Dieses Lernsystem, das sogenannte DECIMER-Lernsystem (Deep IEarning for Chemical ImagE Recognition) kann organisch-chemische Strukturen in digitalen Bildern erkennen. Es wird nicht nur das betreffende Molekül identifiziert, sondern auch alle seine verschiedenen Atome, deren chemische Bindungen, sowie die räumliche Anordnung. Das DECIMER Lernsystem geht über die Naturstoffe hinaus und ist somit eines der leistungsfähigsten OCSR-Systeme (Optical Chemical Structure Recognition). Nun kam die Frage auf, ob das DECIMER-System neben maschinell erstellten Molekülbildern auch handgezeichnete Moleküle erkennen kann.

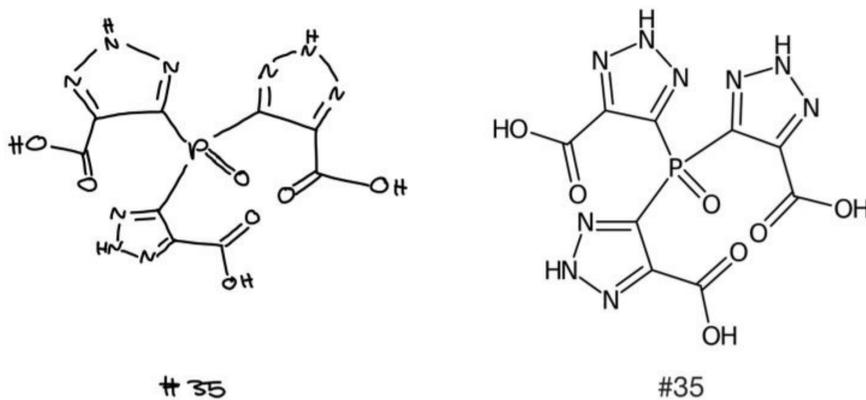


Abb. 1:
Handgezeichnetes und maschinell
erstelltes Molekül

Dies hat das Projekt motiviert, jedoch fehlten noch standardisierte Datensätze mit handgemalten chemischen Strukturen. Hierfür meldeten sich freiwillige Studierende aus dem Studiengang der Molekularbiologie, der Westfälischen Hochschule in Recklinghausen, die unter der Leitung der Jenaer Forscher Dr. Kohulan Rajan und Henning Otto Brinkhaus 5.976 handgemalte Moleküle zeichneten. Aus ca. 100 Millionen chemischen Molekülen war das Ziel 1000 handgezeichnete Moleküle zu erreichen. Mit 5.976 handgemalten Molekülen wurden „die kühnsten Erwartungen des „DECIMER“-Projektteams weit übertroffen“ [1]. Der erreichte Datensatz wird nach wissenschaftlicher Kontrolle auf einer offenen Daten-Plattform publiziert.

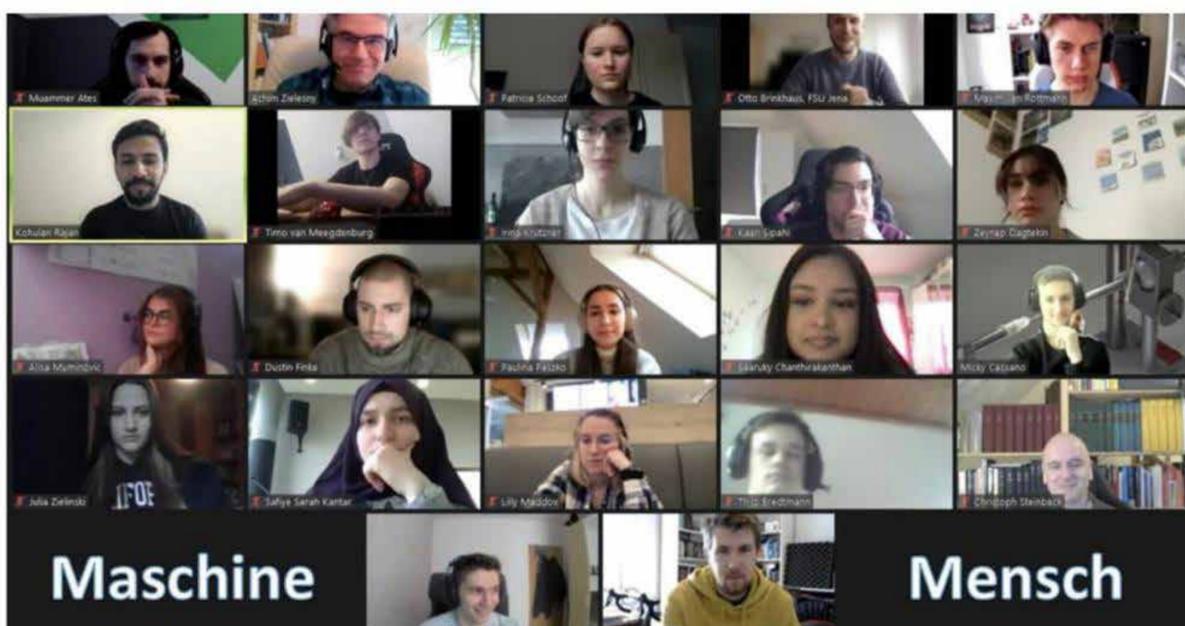


Abb. 2:
Die freiwilligen Studenten und die
Leiter des Projekts

Quelle

[1] Prof. Dr. A. Zielesny, Ausgabe 3/2022, „Handgemalte Moleküle für die künstliche chemische Intelligenz“ aus Trikon Lehre, Seite 6.