

# IBM veröffentlicht Testdaten für KI-gestützte Softwaremodernisierung

[Originalartikel](#)

[Backup](#)

<html> <header class=„article-header“><h1 class=„articleheading“>IBM ver&#246;ffentlicht Testdaten f&#252;r KI-gest&#252;tzte Softwaremodernisierung</h1><div class=„publish-info“> Robert Lippert</div></header><figure class=„aufmacherbild“><img src=„[https://heise.cloudimg.io/width/700/q75.png-lossy-75.webp-lossy-75.foil1/\\_www-heise-de/\\_imgs/18/3/1/0/5/4/1/4/KI-02-9b952b62e091f175.jpeg](https://heise.cloudimg.io/width/700/q75.png-lossy-75.webp-lossy-75.foil1/_www-heise-de/_imgs/18/3/1/0/5/4/1/4/KI-02-9b952b62e091f175.jpeg)“ srcset=„[https://heise.cloudimg.io/width/700/q75.png-lossy-75.webp-lossy-75.foil1/\\_www-heise-de/\\_imgs/18/3/1/0/5/4/1/4/KI-02-9b952b62e091f175.jpeg](https://heise.cloudimg.io/width/700/q75.png-lossy-75.webp-lossy-75.foil1/_www-heise-de/_imgs/18/3/1/0/5/4/1/4/KI-02-9b952b62e091f175.jpeg) 700w, [https://heise.cloudimg.io/width/1050/q75.png-lossy-75.webp-lossy-75.foil1/\\_www-heise-de/\\_imgs/18/3/1/0/5/4/1/4/KI-02-9b952b62e091f175.jpeg](https://heise.cloudimg.io/width/1050/q75.png-lossy-75.webp-lossy-75.foil1/_www-heise-de/_imgs/18/3/1/0/5/4/1/4/KI-02-9b952b62e091f175.jpeg) 1050w, [https://heise.cloudimg.io/width/1500/q75.png-lossy-75.webp-lossy-75.foil1/\\_www-heise-de/\\_imgs/18/3/1/0/5/4/1/4/KI-02-9b952b62e091f175.jpeg](https://heise.cloudimg.io/width/1500/q75.png-lossy-75.webp-lossy-75.foil1/_www-heise-de/_imgs/18/3/1/0/5/4/1/4/KI-02-9b952b62e091f175.jpeg) 1500w, [https://heise.cloudimg.io/width/2300/q75.png-lossy-75.webp-lossy-75.foil1/\\_www-heise-de/\\_imgs/18/3/1/0/5/4/1/4/KI-02-9b952b62e091f175.jpeg](https://heise.cloudimg.io/width/2300/q75.png-lossy-75.webp-lossy-75.foil1/_www-heise-de/_imgs/18/3/1/0/5/4/1/4/KI-02-9b952b62e091f175.jpeg) 2300w“ alt=„IBM ver&#246;ffentlicht Testdaten f&#252;r KI-gest&#252;tzte Modernisierung“ class=„img-responsive“ referrerpolicy=„no-referrer“ /><figcaption class=„akwa-caption“>(Bild:&#160;<a href=„<https://pixabay.com/de/users/geralt-9301/>“ target=„\_blank“ rel=„external noopener“>Gerd Altmann</a>, gemeinfrei)</figcaption></figure><p><strong>Umfangreich und divers: IBMs Project CodeNet liefert Testdaten f&#252;r die Entwicklung k&#252;nstlicher Intelligenzen f&#252;r die Softwaremodernisierung.</strong></p><p>Rund 14 Millionen Codebeispiele und &#252;ber 500 Millionen Zeilen Quellcode zu 55 unterschiedlichen Programmiersprachen &#8211; mit Project CodeNet hat IBM jetzt ein umfangreiches Set an Testdaten f&#252;r die Entwicklung k&#252;nstlicher Intelligenzen f&#252;r die Softwareentwicklung und -pflege ver&#246;ffentlicht. Es soll die Grundlage schaffen f&#252;r automatisierte Code-Korrekturen und Code-&#220;bersetzungen und damit die KI-gest&#252;tzte Modernisierung komplexer, monolithischer Anwendungen erleichtern. Insbesondere sind in dem Set Daten f&#252;r popul&#228;re Sprachen wie C, C++, Java und Python enthalten.</p><p>Als Demonstration f&#252;r die Leistungsf&#228;higkeit des Datensets berufen sich die Forscher bei IBM auf ein nicht n&#228;her benanntes Referenzprojekt, in dem sie einen auf ein Jahr ausgelegten Migrationsprozess mithilfe des KI-Stacks aus CodeNet in nur vier Wochen umsetzen konnten. Dabei wurde eine Legacy-Applikation in &#252;ber 25 Cloud-native Microservices zerlegt.</p><h3 class=„subheading“ id=„nav\_programmiersprac0“>Programmierersprachen schwer zu &#252;bersetzen</h3><p><a href=„<https://www.research.ibm.com/blog/codenet-ai-for-code>“ rel=„external noopener“ target=„\_blank“><strong>In einem Blogbeitrag [1]</strong></a> gehen die Wissenschaftler n&#228;her auf die Herausforderung ein, dass Programmiersprachen sich nicht ohne Weiteres regelbasiert &#252;bersetzen lie&#223;en. Gerade in gro&#223;en Codebasen erschlie&#223;e sich der Kontext einzelner Statements erst &#252;ber mehrere Bibliotheken hinweg. Mit bestehender Technologie lie&#223;en sich dabei nur bis zu 60 Prozent eines Programms &#252;bersetzen; mit ihrem KI-Stack aus CodeNet versprechen die Forscher an dieser Stelle eine deutliche Verbesserung.</p><p>Neben reinen Codebeispielen liefert IBM mit CodeNet auch diverse Metadaten und Annotationen, von Codegr&#246;&#223;e &#252;ber Speicherverbrauch bis hin zu CPU-Laufzeiten, was darauf aufbauenden KIs eine Unterscheidung zwischen Akzeptanz- und Fehlerkriterien erleichtern soll.</p><h3 class=„subheading“ id=„nav\_das\_imagenet1“>Das ImageNet der Softwaremodernisierung?</h3><p>Nach Einsch&#228;tzung von IBM k&#246;nne CodeNet sich

langfristig f&#252;r die Entwicklung k&#252;nstlicher Intelligenzen &#228;hnlich erfolgreich erweisen, wie ImageNet es im Bereich des maschinellen Sehens vorgemacht hat. Die Praxis wird zeigen, ob CodeNet dabei auch eine vergleichbare Fehlerquote aufweist &#8211; erst j&#252;ngst hat eine Studie zweier Wissenschaftler des MIT zusammen mit einem Amazon-Mitarbeiter <a href=„<https://www.heise.de/news/KI-Studie-findet-viele-falsche-Beschreibungen-in-Machine-Learning-Datensaetzen-6000541.html>“><strong>die Zuverl&#228;ssigkeit von Machine-Learning-Datens&#228;tzen untersucht [2]</strong></a>. In dem f&#252;r die Bildklassifikation h&#228;ufig verwendeten ImageNet-Datensatz hat die Pr&#252;fung eine Fehlerquote von 5,83 Prozent ausgemacht.</p><p>Das CodeNet-Datenset <a href=„[https://github.com/IBM/Project\\_CodeNet](https://github.com/IBM/Project_CodeNet)“ rel=„external noopener“ target=„\_blank“><strong>steht auf GitHub [3]</strong></a> unter der freien Apache-Lizenz zur Verf&#252;gung. Begleitend finden sich hier auch Tools unter anderem f&#252;r die statistische Auswertung oder die &#220;bersetzung zwischen beliebten Datenformaten.</p><p>() </p><hr /><p><strong>URL dieses Artikels:</strong><br /><small><code><https://www.heise.de/-6045282></code></small></p><p><strong>Links in diesem Artikel:</strong><br /><small><code><strong>[1]</strong>&#160;<https://www.research.ibm.com/blog/codenet-ai-for-code></code></small><br /><small><code><strong>[2]</strong>&#160;<https://www.heise.de/news/KI-Studie-findet-viele-falsche-Beschreibungen-in-Machine-Learning-Datensaetzen-6000541.html></code></small><br /><small><code><strong>[3]</strong>&#160;[https://github.com/IBM/Project\\_CodeNet](https://github.com/IBM/Project_CodeNet)</code></small><br /><small><code><strong>[4]</strong>&#160;<mailto:map@ix.de></code></small><br /></p><p class=„printversioncopyright“><em>Copyright &#169; 2021 Heise Medien</em></p></html>

From:  
<https://schnipsl.qgelm.de/> - Qgelm

Permanent link:  
<https://schnipsl.qgelm.de/doku.php?id=wallabag:wb2ibm-verffentlicht-testdaten-fr-ki-gesttzte-softwaremodernisierung>

Last update: 2025/06/27 11:17

