

# Mit Forschungsergebnissen die Welt verändern

---

Ein Wissenschaftler sitzt in seinem stillen Kämmerlein und forscht... und dann? Im Interview spricht Dr. Jens Krüger darüber, wie Forschungsergebnisse den Weg in die Unternehmen und von dort aus in den Alltag der Menschen finden. Er ist Fraunhofer-Referent für das Strategische Forschungsfeld »Next Generation Computing«. Das steht auf drei Säulen: die erste Säule basiert auf klassischen Architekturen. Die zweite Säule sind neuromorphe Computer, die in etwa so funktionieren wie unser Gehirn und die dritte Säule sind Quantencomputer.

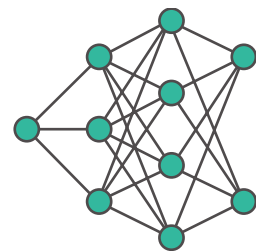
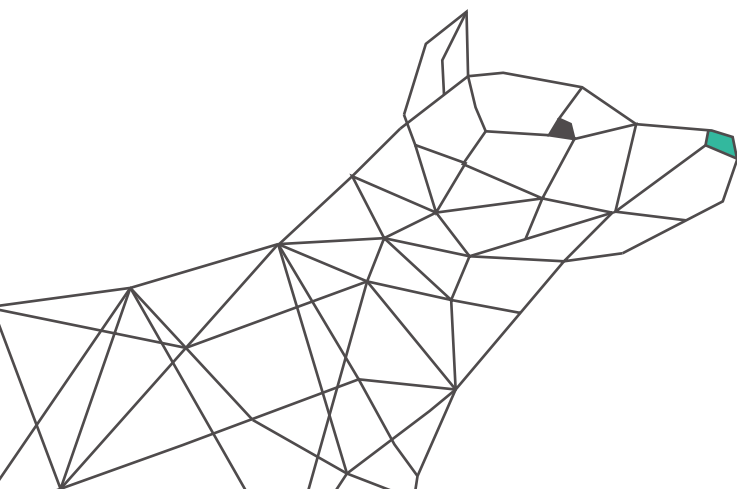
## Fassen Sie bitte kurz zusammen, was Sie als Forscher antreibt?

Ich bin ein neugieriger Mensch. Ich möchte neue Ideen ausprobieren und diese weiterentwickeln bis hin zu Produkten, die dann unserer Gesellschaft und Wirtschaft nutzen. Das geht von der Entwicklung von hocheffizienten Prozessoren bis hin zur Optimierung von künstlichen Neuronen Netzen für mobile Geräte. Ein Beispiel sind Uhren, die ein EKG aufnehmen. Auf diese Weise kann das Gerät einen nahenden Herzinfarkt detektieren und frühzeitig Alarm schlagen. Diese Technologie hat das Potenzial, Menschen das Leben zu retten.

## Oft geben Wettbewerbe der Forschung Impulse. Im März 2021 haben Sie einen Preis im Pilotinnovationswettbewerb »Energieeffiziente KI-Systeme« des Bundesforschungsministeriums (BMBF) gewonnen. Worum ging es da?

Aufgabe war es, eine möglichst energieeffiziente Hardware für den Einsatz in KI-Systemen zu entwickeln, die in EKG-Daten Herzrhythmusstörungen und Vorhofflimmern mit mindestens 90 Prozent Genauigkeit erkennt. Das Projekt, mit dem wir am Wettbewerb teilgenommen haben, hieß HALF, das steht für »Holistischer Ansatz zur Optimierung von FPGA-Architekturen für tiefe neuronale Netze via AutoML – Automatisches Maschinenlernen«. Wir haben dabei die gegenseitige Abhängigkeit des Energieverbrauchs der Hardware und der neuronalen Netzwerktopologie untersucht.

Die Wahl des Netzes hat erheblichen Einfluss auf die Hardwarekomplexität – und damit auf die benötigte Energie und umgekehrt. Diese Abhängigkeiten haben wir optimiert und eine neue Methodik entwickelt, die nicht nur energieeffizienter ist, sondern auch die Entwicklungszeit für optimale neuronale Netzwerktopologien reduziert.



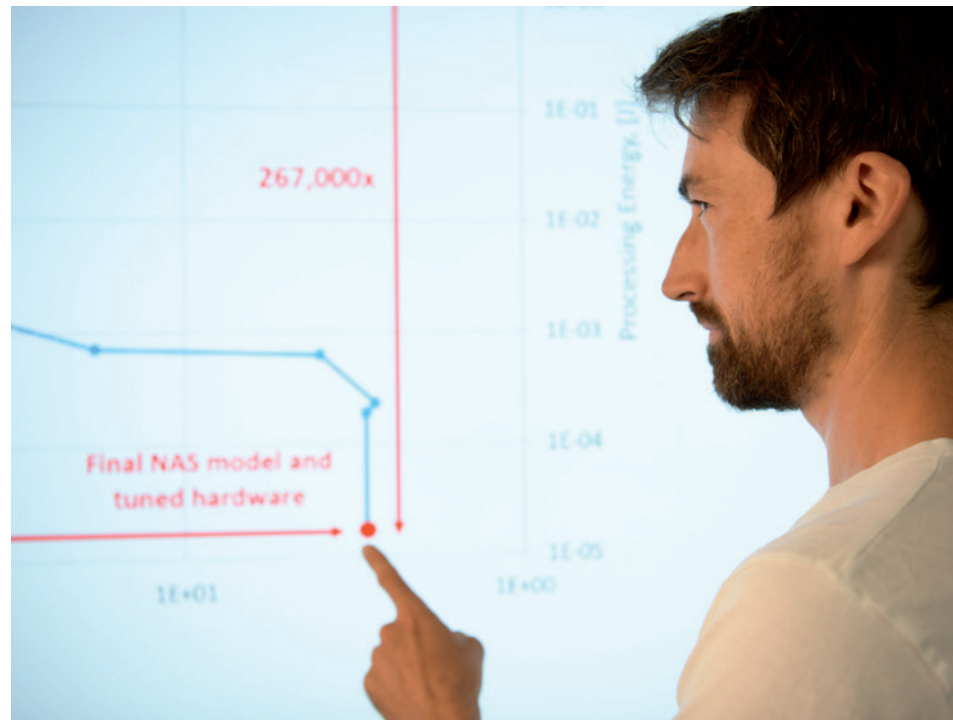
NASE – Neural Architecture Search Engine



## Wie ging es nach dem Gewinn des Wettbewerbs weiter?

Wir stellen nun unsere Expertise Unternehmen zur Verfügung, damit sie ihre Produkte entsprechend weiterentwickeln können. Der KI-Chip im Rahmen des Wettbewerbs war lediglich eine Testplattform. Im Nachfolgeprojekt wird nun direkt mit einem Hersteller zusammengearbeitet, um an der nächsten Gerätegeneration zu arbeiten, welche dann in klinischen Studien eingesetzt werden sollen.

Allerdings können wir alle Branchen unterstützen, denn fast jede steht vor der Herausforderung, dass Datenmengen ständig zunehmen und KI helfen kann, diese zu verarbeiten. Das ist wirtschaftlich für fast alle interessant, zum Beispiel für die Fahrzeugindustrie oder die Telekommunikationsbranche. Daraus entstand das Software-Produkt »NASE« (Neural Architecture Search Engine).



## Mit NASE wird wissenschaftliche Expertise für Unternehmen verfügbar. Wie läuft das ab?

Jeder Auftrag ist individuell, aber klar ist: Effizienz beginnt für uns beim Algorithmus. Wir nutzen modernste Verfahren der automatischen neuronalen Netzwerksuche, um Netzwerke zu entwickeln, die hinsichtlich vieler Aspekte gleichzeitig effizient sein können. Wir beachten Eigenheiten der zugrundeliegenden Plattform und nehmen sie in das Netzwerkdesign auf. Der Algorithmus passt dann die Netze der Hardware an. Wir bieten also unsere Erfahrung, die Technologie und Rechenkapazitäten. Die Unternehmen liefern uns die für sie relevanten Datensätze und definieren die Anforderungen, etwa an Genauigkeit und Geschwindigkeit. Dann suchen wir mithilfe unserer Supercomputer und unseres Frameworks nach dem besten Modell. Das Netzwerk ist dann direkt einsatzfähig.

**Um dem Bedarf nach immer mehr und schnellerer Rechenleistung gerecht zu werden, forschen Sie in der European Processor Initiative (EPI) gemeinsam**

## mit 28 Partnern aus 10 europäischen Ländern und entwickeln hocheffiziente Beschleunigerprozessoren. Welchen Beitrag leistet das Fraunhofer ITWM hier?

Unser Beitrag ist der sogenannte Stencil- und Tensor Beschleuniger (STX), den wir gemeinsam mit dem Fraunhofer IIS auf Basis einer Architektur der ETH Zürich entwickeln. Wir konzentrieren uns auf das effiziente Ausführen von hoch parallelisierbaren Anwendungen mit konstanten Zugriffsmustern, wie sie in vielen Anwendungen vorkommen – von der Strömungsdynamik, über Klima- und Wettervorhersagen bis hin zu bildgebenden Verfahren. Reale Anwendungen sollen energieeffizienter werden, einfacher zu programmieren sein und geringere Kosten verursachen. Schon jetzt können Interessenten eigene Codes auf unserem Simulator testen. Nächstes Jahr steht dann die nächste Generation von Testchips zur Verfügung. 2025 wollen wir das erste komplette System lauffähig haben. Eine große Herausforderung, aber auch ein wichtiger Schritt hin zu einer neuen nationalen und europäischen Industrie für hochperformante Prozessoren und Beschleuniger.

### Kontakt

Dr. Jens Krüger  
Sprecher der Leitung des Bereichs  
»High Performance Computing«  
Telefon +49 631 31600-4541  
jens.krueger@itwm.fraunhofer.de

