



Chemie . Festkörperforschung . Materialwissenschaften

## „Die Industrie hat großes Interesse an einer Enzyklopädie der Werkstoffe“

Interview mit Matthias Scheffler zum europäischen Exzellenzzentrum NOMAD

2. November 2015

**Das Europäische Exzellenzzentrum NOMAD (Novel Materials Discovery – Entdeckung neuartiger Materialien) soll die Suche nach geeigneten Materialien für neue technische Entwicklungen und die akademische Forschung beschleunigen. Diesem Ziel haben sich acht Forschungseinrichtungen, darunter zwei Max-Planck-Institute, und vier Hochleistungsrechenzentren wie etwa die Max Planck Computing and Data Facility in Garching aus ganz Europa verschrieben. Zum offiziellen Start des Exzellenzzentrums am 1. November 2015 sprachen wir mit Matthias Scheffler, Direktor am Fritz-Haber-Institut der Max-Planck-Gesellschaft und Leiter des Forschungsverbundes, über dessen Ziele und Besonderheiten.**



Matthias Scheffler, Direktor am Fritz-Haber-Institut der Max-Planck-Gesellschaft und Leiter des europäischen Exzellenzzentrums NOMAD. © Steffen Kangowski

*Worin liegt die Motivation für die Gründung des Exzellenzzentrums NOMAD?*

Um neue Produkte zu entwickeln, ist die Industrie fast immer auf neue Werkstoffe angewiesen. Nehmen Sie zum Beispiel ein Smartphone. Vor 20 Jahren gab es noch keine Materialien für zuverlässige und kompakte, berührungsempfindliche Bildschirme. Und auch noch keine Halbleiter für ausreichend leistungsfähige Elektronik. Erst die Fortschritte in der

Werkstoffwissenschaft haben Smartphones, die sich per Touchscreen bedienen lassen, möglich gemacht. Die industrielle Bedeutung der Materialwissenschaften hat auch das Weiße Haus erkannt, und US-Präsident Obama hat im Juni 2011 die *Materials Genome Initiative for Global Competitiveness* angekündigt. Die Anspielung auf das Human-Genom-Projekt im Namen der Initiative illustriert die umfassende und forcierte Herangehensweise. Seit Ende 2013 gibt es für das Screening von Werkstoffen im Hochdurchsatzverfahren, um das es in der Initiative geht, in den USA massive Forschungsmittel.

*Welche Ziele verfolgen Sie mit NOMAD?*

In unserem Exzellenzzentrum wollen wir neue Methoden entwickeln, mit denen die riesigen Datenmengen der Materialwissenschaften, die bereits bestehen und die auch wir täglich vergrößern, systematisch analysiert werden können. Dies wird es ermöglichen, noch unbekannte Materialien mit interessanten Eigenschaften zu finden und in Materialien, die wir bereits kennen, bislang unentdeckte Phänomene aufzuspüren. Letztlich wollen wir auch eine virtuelle Enzyklopädie der Materialien aufbauen.

*Worin unterscheidet sich Ihr Exzellenzzentrum von der Materials Genome Initiative?*

In der Initiative der US-Regierung ging es bislang vornehmlich darum, möglichst viele fundamentale Daten von Materialien und ihren Eigenschaften in Berechnungen und experimentellen Untersuchungen zu erzeugen. Die *Materials Genome Initiative* geht davon aus, dass man in diesen Daten die benötigten Informationen direkt finden kann. Da bräuchte man aber einen beinahe unendlich großen Datensatz, weil der Raum möglicher Werkstoffe, die wissenschaftlich und technologisch interessant sein können, praktisch unendlich groß ist. Hierbei denke ich an stabile, aber auch an metastabile Materialien sowie an Hybridsysteme und Nanostrukturen. In unserem Exzellenzzentrum werden wir uns zunächst auf berechnete Daten konzentrieren und wie bisher, auch weitere Daten berechnen. Im Wesentlichen wollen wir aber einen Schritt weiter gehen, und die Analyse der großen Datenmengen der Materialwissenschaften vorantreiben.

*Das heißt?*

Einfach formuliert: Die Berechnung riesiger Datenmengen auf den größten Supercomputern gewinnt erst dann großen Wert, wenn wir auch die entsprechenden *big data analytics* Methoden haben. Diese fehlen aber in den Materialwissenschaften fast völlig, und die Entwicklung ist äußerst kompliziert. Der Schwerpunkt unserer Entwicklung liegt daher auf Konzepten und Methoden der angewandten Mathematik. Ein Stichwort ist etwa *compressed sensing*. Dieser Ansatz ist noch keine zehn Jahre alt und wurde in den Materialwissenschaften bislang noch nicht wirklich verfolgt. Mit diesem Verfahren wollen wir beispielsweise Muster in den *big data* der Materialien aufspüren, die man bisher nicht erkennen konnte. Damit kann man dann auch bislang unbekannte Materialien und ihre Eigenschaften wesentlich effizienter vorhersagen.

*Können Sie die Methoden kurz erläutern?*

Rechnergestützte Materialwissenschaften sind der Kern des Exzellenzzentrums: Wir lösen die Vielteilchen-Quantenmechanik von Materialien mit Hilfe von sehr

komplexen Programmpaketen auf Großrechnern sehr genau, um die Eigenschaften von Materialien zu beschreiben und Funktionen, für die ein Material verwendet werden kann, zu identifizieren. Die so berechneten Daten von mehreren hunderttausend Materialien müssen dann weiter analysiert werden, zum Beispiel mit compressed sensing.

*Worum geht es dabei?*

Das Verfahren wurde in den Computer- und Ingenieurwissenschaften für die Signalverarbeitung entwickelt. Grundlage des Verfahrens ist, dass große Datensätze oft eine interne Struktur besitzen, so dass sie mit einer relativ kleinen, kompakten Darstellung beschrieben werden können. Es zeigt sich, dass diese dünne Verteilung – im Englischen heißt das *sparsity* – eher die Regel ist als die Ausnahme. Allerdings ist zunächst unbekannt, wie man die Daten darstellen muss, um diese *sparsity* zu identifizieren. Die Methoden für das Auffinden solcher Datenstrukturen wollen wir nun weiterentwickeln, um sie auf die Materialwissenschaften zuzuschneiden.

*Was bringt die kompakte Darstellung der Daten?*

Die Kenntnis von Strukturen in großen Datenmengen und die dadurch mögliche Reduktion auf das Wesentliche erlaubt dann die effiziente Vorhersage von bislang nicht berechneten Materialien und ihren Eigenschaften. Wir stehen hier noch ganz am Anfang. Aber immerhin haben sich in NOMAD die führenden Festkörpertheoretiker Europas zusammengeschlossen, um in dieser Richtung voranzukommen.

*Sie setzen bei NOMAD ausschließlich auf Daten theoretischer Untersuchungen, also auf Rechnungen. Warum verwenden Sie keine Ergebnisse experimenteller Messungen?*

Der Hauptgrund ist, dass wir zunächst die erwähnten Verfahren entwickeln oder weiterentwickeln müssen. Es gibt bislang nur sehr wenige Anwendungen der *big data analytics* in den Materialwissenschaften. Für die geplanten Entwicklungen ist es deshalb wichtig, die Fehler oder Ungenauigkeiten der Daten zu kennen, und beim Hochdurchsatz-Screening im experimentellen Bereich ist das nicht unbedingt der Fall. Zu einem späteren Zeitpunkt werden wir aber auch experimentelle Daten berücksichtigen.

*Welche Rolle spielt die Industrie bei NOMAD?*

Wir haben ein starkes Beratungsgremium aus der Industrie, die an einer Enzyklopädie der Werkstoffe und an den erwähnten Analysemethoden großes Interesse hat. Wir werden uns regelmäßig mit Kollegen verschiedener Unternehmen treffen. Dabei wollen wir von den Managern die grundsätzlichen Erfordernisse der verschiedenen Industrieunternehmen erfahren und mit aktiven Wissenschaftlern aus der Industrie werden wir gemeinsam an konkreten Fragestellungen arbeiten. Problematisch wird eine konkrete Zusammenarbeit allerdings dann, wenn wir die Arbeiten nicht öffentlich machen können. Wir werden deshalb auch Trainingskurse für die Kollegen aus der Industrie anbieten, um die von uns entwickelten Methoden weiterzugeben. Dann kann die Industrie die Methoden auch ohne unser Zutun verwenden.

*Wie schwierig war es, die Fachwelt davon zu überzeugen, die Enzyklopädie von NOMAD und alle Daten, auf denen sie beruht, frei zugänglich zu machen?*

Vor zwei Jahren gab es da noch große Widerstände. Viele Kollegen, aber vor allem auch Vertreter der Industrie hatten Bedenken, dass dann auch Forscher außerhalb Europas, etwa in Fernost, über die Daten verfügen können. Ich meine aber, dass die Wissenschaft ihre Erkenntnisse frei zugänglich machen sollte. Einige Kollegen waren und sind zurückhaltend ihre Daten vollständig zur Verfügung zu stellen, weil die Berechnung ja eine Menge Arbeit gemacht hat. Letztlich werden durch NOMAD auch Wissenschaftler und Ingenieure die Daten verwenden können, die kein Wissen über die zugrunde liegenden komplexen Methoden haben. Sie werden eventuell Verfahren verwenden, von denen wir noch nichts wissen. Ich erwarte viele Überraschungen, was Leute mit den Daten machen, wenn sie verfügbar gemacht werden, und ich bin überzeugt, dass dies die Wissenschaft nachhaltig voranbringen wird. Mehr und mehr Kollegen schließen sich dieser Auffassung an, weshalb wir schon jetzt Ergebnisse von fast einer Million Rechnungen in unserem NOMAD-Repository haben.

*Können Sie schon abschätzen, ab wann die Enzyklopädie von NOMAD online sein wird?*

In den letzten Monaten haben wir gemeinsam mit Claudia Draxl von der Humboldt-Universität und dem Hochleistungsrechenzentrum der Max-Planck-Gesellschaft in Garching die Datenbank NOMAD-Repository.eu aufgebaut. Der Umfang der Daten, die heute dort liegen, ist schon jetzt immens. Der Aufbau der NOMAD-Datenbank, in der diese Daten auf ein einheitliches Format gebracht werden und dann für die *big data analytics* geeignet vorliegen werden, wird wohl sechs bis acht Monate dauern. Eine erste Version der Enzyklopädie der Materialwissenschaften sollte in eineinhalb bis zwei Jahren vorliegen.

*Das Interview führte Peter Hergersberg*