

# Digitalgespräch Folge 23

## Hochleistungsrechnen zu Zukunftsfragen: Das Deutsche Klimarechenzentrum

Mit Thomas Ludwig vom Deutschen Klimarechenzentrum, 14. Juni 2022

<https://zevedi.de/digitalgespraech-023-thomas-ludwig/>

*[Der Vorspann mit Musik und Ausschnitten aus dem Gespräch beginnt.]*

**Marlene Görger [mg]:** Herr Ludwig, Sie sind Professor am Fachbereich für Informatik der Universität Hamburg und Direktor des Deutschen Klimarechenzentrums.

**Thomas Ludwig [Ludwig]:** Die Modelle zu dieser Zeit hatten überhaupt keine Mechanismen für Energieeffizienz. Damals war das einfach so: Ja, naja, Hochleistungsrechnen, das ist so die Formel 1 des Rechnens und es braucht halt mehr und so ist es.

**Petra Gehring [pgg]:** Kann man sich das so vorstellen, dass die Bekämpfung des Klimawandels tatsächlich alle internationalen Akteure friedlich und kollaborativ zusammenbringt?

**[Ludwig]:** Nehmen wir mal an, wir könnten auf einmal einen Rechner hinstellen, der Faktor 1000 schneller ist als der vorhandene. Dann würden die Wissenschaftler aktuell gar nicht wissen, wie sie diese Leistung ausnutzen sollen.

**[mg]:** Was verspricht man sich denn von maschinellem Lernen?

**[Ludwig]:** Das wird eine interessante Sache, weil da habe ich eine Meinung als DKZ-Chef und eine Meinung als Wissenschaftler. Schlecht in Mathe zu sein, ist in Deutschland irgendwie schick. Ungünstig für eine moderne Gesellschaft, bei der so viel auf IT dann basiert. Dieses Nicht-Wahrnehmen, dass diese Infrastruktur halt sehr wichtig ist.

*[Der Vorspann endet, das Gespräch beginnt.]*

**[mg]:** Der vom Menschen verursachte Klimawandel wurde schon seit über 100 Jahren von Wissenschaftlern vorhergesagt. Bereits 1906 hat der schwedische Physiker und Chemiker Svante Arrhenius Prognosen zur Erderwärmung infolge der Verbrennung von Kohle veröffentlicht. Die Berechnungen dafür musste und konnte er damals auch ohne Supercomputer durchführen. Nun hat der Klimawandel längst begonnen, viel schneller und heftiger als Arrhenius es für möglich gehalten hätte. Denn seit seiner Zeit haben sich die globalen CO<sub>2</sub>-Emissionen um etliche Größenordnungen erhöht. Und Modelle, mit denen die tatsächlichen Entwicklungen beschrieben werden können, sind seit den Anfängen der Klimaforschung wesentlich komplexer geworden. Sie zu beherrschen und zu benutzen, wäre ohne moderne Supercomputer unverstellbar. Für die Verarbeitung gigantischer Datensätze in immer aufwendigeren Berechnungen kommen modernste Höchstleistungsrechner zum Einsatz. Hierzulande geschieht dies in Hamburg. Dort gibt es seit 1987 das Deutsche Klimarechenzentrum. 1988 nahm es die Arbeit auf und legte mit der Anschaffung seines ersten Supercomputers einen wichtigen Grundstein für die Klimaforschung in Deutschland. Der erste Direktor des DKRZ, Klaus Hasselmann, wurde für seine Arbeit zu

Klimamodellen im Jahr 2021 mit dem Nobelpreis für Physik ausgezeichnet. Wie hat sich das Hochleistungsrechnen für die Klimaforschung seitdem entwickelt? Wie ist das heute organisiert? Welche Möglichkeiten bietet es, jetzt wie auch in Zukunft? Und wie sieht die Ökobilanz von Rechenzentren aus? Muss sie verbessert werden? Und wenn ja, wie könnte das gehen? Auch darüber wollen wir heute im Digitalgespräch reden. Mein Name ist Marlene Görger. Ich bin Physikerin und Technikphilosophin am Zentrum verantwortungsbewusste Digitalisierung.

**[pgg]:** Und ich bin Petra Gehring, Professorin für Philosophie an der TU Darmstadt. Wir haben wie immer einen Gast hier bei uns im Digitalgespräch, einen Experten für wissenschaftliches Hochleistungsrechnen, nämlich Prof. Dr. Thomas Ludwig aus Hamburg. Herzlich willkommen im ZEVEDI-Podcast, Herr Ludwig. Vielen Dank, dass Sie sich heute Zeit für uns nehmen.

**[Ludwig]:** Gerne.

**[mg]:** Herr Ludwig, Sie sind Professor am Fachbereich für Informatik der Universität Hamburg und Direktor des Deutschen Klimarechenzentrums. Ihr Fachgebiet ist das wissenschaftliche Rechnen. Und das ist natürlich auch, was im Klimarechenzentrum passiert. Das DKRZ ist, so heißt es auf der Webseite, praktisch das Labor aller deutschen Klimaforscher, die ihre Ergebnisse auf der Basis der Modellierung des Klimas erzielen. Wie kam es denn, dass das DKRZ diese zentrale Funktion hat? War das von vornherein so geplant?

**[Ludwig]:** Also das DKZ wurde ja 1987 aus der Taufe gehoben. Im Januar 88 ging es dann in den Operationsbetrieb über. Ich habe gerade kürzlich erst festgestellt, dass wir ja demnächst 35-jähriges Jubiläum haben, was ganz schön ist. Und es ist als Organisationseinheit entstanden aus einem gemeinschaftlichen Rechenzentrum, das die Universität Hamburg hatte, zusammen mit dem Max-Planck-Institut für Meteorologie. Max-Planck-Institut für Meteorologie ist hier im Nachbarhaus angesiedelt. Die Universität hat im Nachbarhaus ihr Institut für Geowissenschaften, die zusammen hatten einen Rechner betrieben. Und man hatte dann den Plan, eine unabhängige Einheit zu gründen. Es war ein bisschen auch damit verbunden, dass das Ministerium damals eingestiegen ist in die Förderung eines Rechners. Und es erschien günstig, das als eine GmbH zu betreiben. In dem Fall: Wir sind jetzt eine gemeinnützige GmbH, verwalten aber öffentliche Gelder. Im Prinzip funktionieren wir eigentlich wie ein normales Forschungsinstitut, nur dass wir eben von der Rechtsform her als Firma betrieben werden.

**[pgg]:** Also was man ja quasi schon am Namen hört, ist, dass das DKRZ ein Dienstleister für Forschende ist, die sich auf Klimaforschung spezialisiert haben. Aber gibt es jenseits, sagen wir mal, dieses Kundenkreises auch noch handfeste Spezifikationen der Aufgabenstellung? Sehen die Rechner anders aus oder werden die anders benutzt als für andere Themen, für andere Forschungsfragen?

**[Ludwig]:** Vielleicht bleiben wir zuerst mal bei den Kunden selber. Dieses Konstrukt mit dem DKZ als gemeinnütziger GmbH erfordert, dass man irgendwie eine Satzung hat mit einem Satzungszweck. Und bei uns ist dieser Zweck die Förderung der Forschung in den Erdsystemwissenschaften. Die Begrifflichkeit ist immer so ein bisschen schwierig, ob man das Klimaforschung nennt, das ist meistens zu eng. Erdsystem ist meistens etwas zu weit, weil in dem Zusammenhang zum Beispiel auch

Plattentektonik und Erdbebenforscher eine Rolle spielen, die bei uns gar nicht rechnen. Die haben zum Teil gar keine Modelle, die sie da verwenden könnten. Also insofern haben wir eine satzungsgemäße Bindung an einem bestimmten Personenkreis, die hier jetzt ihre Forschung betreiben. Und es ist schon so, dass die Rechner, die für diese spezielle Community betrieben werden, typischerweise ein bisschen anders konfiguriert sind, als wenn man ein Hochleistungsrechenzentrum nimmt für, sagen wir mal, eine Vielzahl von Communities. Die drei deutschen sogenannten Höchstleistungsrechenzentren – Jülich, München und Stuttgart zum Beispiel –, die bedienen verschiedenste Wissenschaftsgemeinschaften und haben Rechner, die halt einfach für alle von denen passend sind. Und bei uns ist das ein bisschen anders konstruiert. Und dann kann man sagen, dass insbesondere beim Klima die Ausstattung mit Speicher, sowohl jetzt die drehenden Festplatten als auch im Hintergrund die Bandarchive, dass das eben sehr üppig ist. Da haben wir in Relation zu unserer Rechenleistung ein Vielfaches von dem, was andere Rechenzentren haben. Und wir haben auch meistens bei der Prozessorarchitektur so ein bisschen speziellere Anforderungen. Ich will mal sagen, bei den Prozessoren hat es damit zu tun: Die Anwendungen, die im Klima gerechnet werden, werden schon sehr lange entwickelt. Und die werden von großen Teams entwickelt. Und die Entwicklungszeiten sind häufig im Bereich von 10 Jahren, 20 Jahren. Das heißt, die Codes sind über lange Zeit gewachsen. Und wenn jetzt eine neue Architektur auf den Markt kommt, dann kann man nicht auf die Schnelle die Codes an diese Architektur anpassen. Also wir sind da immer relativ konservativ aufgestellt. An neuen Architekturen, da gibt es so ein Konstrukt, das nennt sich Grafikkarten. Das sind die Karten, die man im normalen Rechner hat, mit denen eigentlich das Bild generiert wird. Die eignen sich ganz toll für bestimmte Rechenoperationen. Die würden sich auch im Klima gut eignen. Nur sind sie schwerer zu programmieren. Wenn wir die haben wollten, müssten wir die ganzen Codes umschreiben. Aufgrund der Größe der Codes und ihrer langen Entwicklungszeit, die Entwickler sind dann auch gar nicht mehr greifbar in vielen Fällen, geht sowas meistens nicht. Und beim Speichersystem ist es so, da muss man immer sehen: Wenn man Klimamodellierungen macht, so wie bei uns im Haus, dann sind die Berechnungen über einen simulierten Zeitraum von zum Beispiel 50 Jahren, 80 Jahren, 100 Jahren. Und da will man die Ergebnisdaten natürlich schon eine Zeit lang aufheben können. Sagen wir mal vielleicht 10, 20 Jahre, 10 Jahre sicher mindestens, um zu einem späteren Zeitpunkt auch noch mal zu gucken: Ja, ist es denn jetzt auch so gekommen, wie wir das modelliert hatten? Deswegen sind in unserer Community die Anforderungen an Speicherzeit wesentlich höher. Und das ist der Grund, warum wir diese großen Bandarchive betreiben. Also vom Rechner her, wenn wir einen neuen Rechner installieren, sind wir in den Rängen 30 bis 50 vielleicht, dieser Top-500-Liste. Aber was unser Bandarchiv angeht, sind wir sicher bei den fünf größten Speicherzentren weltweit anzusiedeln.

**[pgg]:** Das heißt, diese 35 Jahre, die es das Zentrum schon gibt, und die 100, die es nach vorne rechnet, die erzwingen ein hohes Maß an Nachhaltigkeit? Es hat eigentlich ein sehr auf langfristige Nutzung angelegtes Maß, was passiert am Zentrum?

**[Ludwig]:** Ja, ich vergleiche das immer so ein bisschen mit dem, wie das beim Wetter ist. Bei uns wird ja überhaupt kein Wetter gerechnet. Wetter wird gerechnet beim Deutschen Wetterdienst in Offenbach. Beim Wetter ist es im Prinzip so: Die berechnen diese Vorhersage für die nächsten 14 Tage, sagen wir mal. Und das ist jetzt natürlich ein bisschen stark simplifizierend, aber wenn die 14 Tage rum ist, können sie eigentlich die Daten wegschmeißen, weil, niemand ist an falschen Daten einer alten

Wettervorhersage interessiert. Stimmt nicht ganz, weil natürlich aus der Differenz vom Vorhergesagten und Eingetretenen verbessern die selbstverständlich ihre Wettervorhersagen. Also, auch die heben Daten auf. Aber es ist ein methodisch anderes Konzept als bei uns. Bei uns ist tatsächlich wegen der Langfristigkeit der Vorhersagen auch die langfristige Speicherung eine notwendige Konsequenz.

**[mg]:** Man sieht ja, wenn man sich mal anschaut, welche Rechnertypen über die Jahre bei Ihnen angeschafft wurden, ziemlich deutlich diese enorme Steigerung in der Hardware, was da möglich ist, an Speicher-, an Prozessorleistung. Haben Sie so Beispiele im Kopf für vielleicht auch Sprünge in der Forschung, die möglich wurden durch die Anschaffung neuer Rechner? Gab es da so Momente, wo man wusste, ok, jetzt können wir das endlich machen, was wir vorhatten, und das war wirklich ein Meilenstein?

**[Ludwig]:** Man kann es vielleicht ein bisschen auf folgende Weise illustrieren. Also zwischen den Rechnern, mit denen wir begonnen haben, die im Vergleich zu heute vielleicht die Leistungsfähigkeit von einem wirklich Billig-Smartphone haben, und dem Rechner, den wir jetzt betreiben, haben wir Sprünge, Leistungssteigerungen in der Kategorie von einer zweistelligen Millionenzahl. Ich glaube, innerhalb der ersten 25 Jahre (und jetzt sind wir ja, wie gesagt, bei 35 fast) hatten wir etwa eine Million als Leistungssteigerung, als Rechenleistungssteigerung. Und Volumenspeicherung war in etwa konsistent damit. Und wenn man dann guckt, was über die Jahre hinweg simuliert worden ist, dann war es eben so: Anfänglich haben die in der Klimaforschung beispielsweise nur Land und die Atmosphäre simuliert. Dann kamen so Dinge dazu, dass man den Ozean mit dazu genommen hat. Im ersten Schritt zum Beispiel nur die Ozeanoberfläche. Aber als dann die Rechenleistung größer wurde, hat man zum Beispiel das Innere vom Ozean mit den Strömungen dazu genommen, was ein wichtiger Punkt ist. Als die Rechnerleistung über die Jahrzehnte eben gestiegen ist, ziemlich schnell in der Vergangenheit natürlich, kamen so Dinge hinzu, dass man zum Beispiel diese ganze Kryosphäre mitberechnet hat, alles, was mit Eis zu tun hatte. Also sowohl das Eis, das auf dem Ozean schwimmt, als auch das Landeis, das in den Gletschern gebunden ist. Dann kam hinzu, dass man chemische Prozesse, biologische Prozesse mit aufgenommen hat. Und das war immer so ein Wechselspiel zwischen dem: Was kann man schon modellieren und wofür hat man dann die Rechen- und Speicherleistung, um es tatsächlich auf einem globalen Maßstab auch umsetzen zu können? Es ist so, dass wir uns da in einem Wechselspiel zwischen Machbarkeit in der Wissenschaft und Ausnutzung dieser Rechner befinden. Also manche Dinge sind, sagen wir mal, zehn Jahre im Voraus durchaus denkbar und man hat die Hardware nicht dafür. Aber es ist jetzt auch so, nehmen wir mal an, es gäbe irgendeinen glücklichen Umstand. Wir könnten auf einmal einen Rechner hinstellen, der Faktor 1000 schneller ist als der vorhandene Rechner. Dann würden die Wissenschaftler aktuell gar nicht wissen, wie sie diese Leistung ausnutzen sollen, weil eine Modellbildung, die jetzt diese Leistung abfragt, in ihren Köpfen und auf ihren Zetteln mit der Mathematik noch gar nicht existiert. Momentan befinden wir uns gerade an so einer Schwelle, wo Klima- und Wettermodellierung etwas zusammenkommen, wenn es insbesondere darum geht, Wolken und Niederschlag zu modellieren. Also Wolken und Niederschlag ist ja klar: Das ist im zeitlichen Rahmen ein kurzfristigerer Effekt. An dem sind eigentlich die Leute interessiert, die die Wettervorhersage machen. Aber wenn man in der Lage wäre, das in eine Klimavorhersage richtig hineinzurechnen, wäre viel gewonnen. Zum Beispiel in Bezug auf die Wolken. Da wird bei den Klimamodellierungen quasi nur so eine Art Bewölktheitsparameter mit eingebaut.

Also, der Globus wird ja in so Zellen unterteilt, und dann wird für jede Zelle halt ein Bewölktheitsgrad, ein durchschnittlicher Bewölktheitsgrad, angenommen, in der Sahara wenig und an anderen Stellen eben viel, je nachdem, wie viele Wolken das da so hat. Und man will weg von dieser sogenannten Parametrisierung hin zu einer realen Berechnung. Und für eine reale Berechnung des Wolken- und Niederschlagsgeschehens muss man aber, weil das auch räumlich so kleinteilige Ereignisse sind, bei diesem Gitter, mit dem man es räumlich auflöst, runter auf vielleicht einen Kilometer Abstand. Das würde bedeuten, dass wir Rechner brauchen, die im Vergleich zu jetzt, naja, vielleicht schon diesen Faktor 1000 mindestens schneller sind. Da hat man jetzt aktuell eine Situation, wo man wahrscheinlich, wenn so ein leistungsfähiger Rechner vom Himmel fallen würde, ihn an dieser Stelle tatsächlich schon nutzen könnte. Aber in anderen Bereichen der Fragestellungen wäre das noch nicht so.

**[pgg]:** Das klingt so, als ob das, was gerade die Modellierer und Modelliererinnen machen, sehr eng verbunden ist mit dem, was das Rechenzentrum beiträgt zum Ganzen. Wie sieht denn die Arbeitsteiligkeit aus? Also, es gibt sicherlich diejenigen, die einfach rechnen wollen, die verwenden vorhandene Modelle. Es gibt diejenigen, die die Modelle verändern, verbessern, manipulieren. Es gibt diejenigen, die dann die Modelle auf den Rechner bringen. Wahrscheinlich ist das jetzt eine Klein-Erna-Vorstellung und es gibt noch viel mehr. Aber Sie klingen so, als ob Sie eigentlich alles im Blick hätten.

**[Ludwig]:** Ja, ich denke, wir müssen alles im Blick haben. Und es ist genau wie Sie sagen, es gibt tatsächlich diese verschiedenen Nutzergruppen. Zum Beispiel gibt es, sagen wir mal, Doktorandinnen und Doktoranden, die vorhandene Modelle nutzen, um irgendetwas Bestimmtes auszurechnen oder zu modellieren. Und dann gibt es die Gruppen, meistens sind es Forscherinnen und Forscher, die längerfristig dann an Instituten arbeiten, die tatsächlich auch die Modell- entweder Entwicklung oder Weiterentwicklung vorantreiben. Man muss immer so ein bisschen sehen: Diese Forscherinnen und Forscher, die sind ja an Wahrheit ein Stück weit interessiert. Das ist der Unterschied zu den Wetterfröschen, sage ich immer, weil, die wollen einfach nur schnell mal irgendeine Vorhersage und das soll einigermaßen stimmen. Deswegen können die auch so Dinge machen mit diesem maschinellen Lernen zum Beispiel, wo man hintennach nicht fragen kann: Warum stimmt das jetzt eigentlich? Aber die, die die Klimamodellierung machen, die sind ja quasi an Wahrheit interessiert. Und da gibt es jetzt einen Personenkreis, die diese Programme, und die Programme sind sozusagen die softwaremäßige Repräsentation der Modelle, die also diese Programme vorantreiben. Das typische Vorgehen ist auch, dass es keinen Endzustand von so einem Programm gibt, weil die entwickeln was, die testen, ob das eigentlich von den Ergebnissen her valide ist, also ob das mit dem Beobachteten übereinstimmt, zum Beispiel. Wenn das der Fall ist, freuen sie sich und suchen sich ein neues Ziel. Also Berechnungen zu machen im Sinne eines Produktionsbetriebes, um zum Beispiel zu sagen: In 50 Jahren ist dieses und jenes passiert oder wird passieren, das ist eine andere Kategorie der Nutzung dieser Modelle, als zu sagen: Ich gucke jetzt mal, wie das mit den Wolken ist, und wenn ich das mit den Wolken weiß, dann mache ich noch den Hagel dazu. Und wenn ich es mit dem Hagel weiß, dann fällt mir sicher noch was Neues ein. Insofern haben wir ganz unterschiedliche Gruppen von Menschen, die mit diesen Codes arbeiten. Und die, die das weiterentwickeln, unterstützen wir dabei, dass die auch Kenntnis haben, beispielsweise über neue Rechnerarchitekturen, also wenn auf dem Markt etwas auftaucht, was die Ausführung ihrer Anwendungen beschleunigen könnte, zum Beispiel. Und die andere Gruppe, die unterstützen wir

hauptsächlich durch Serviceleistung, wenn die im operationellen Betrieb Schwierigkeiten haben, eben auch dann die Datenmengen, die die produzieren, vernünftig zu speichern, zu annotieren, wenn eine Visualisierung gemacht wird, ja die Filmchen, die wir dann teilweise im Fernsehen auch sehen können, solche Dinge. Also da haben wir ein sehr breites Portfolio, das alle diese Wünsche dann abdeckt.

**[mg]:** Sie sagten ja, die Rechenleistung und die Möglichkeit, die das Rechenzentrum bietet, haben auch Einfluss darauf, welche Forschungsfragen entwickelt werden können. Vor allen Dingen werden ja die Modelle dann auch komplexer, es werden mehr Faktoren berücksichtigt, mehr Randbedingungen berücksichtigt. Ich bin auf die Bemerkung gestoßen, dass mittlerweile auch sozioökonomische Faktoren eine Rolle spielen können in der Klimamodellierung. Wer ist denn die Gruppe, die solche Fragestellungen dann entwickelt? Da werden ja eigentlich zwei wissenschaftliche Sphären zusammengebracht. Ist das eine neue Entwicklung oder gibt es da schon Ergebnisse? Wissen Sie das?

**[Ludwig]:** Also ich glaube, dass eine Modellierung der sozioökonomischen Gegebenheiten nicht in Form von mathematischen Modellen aktuell eingebunden ist. Oder wenn, dann handelt es sich um Modelle, die nicht auf Hochleistungsrechnern laufen, sondern in nachgeschalteten Vorgängen irgendwie damit verwoben werden. Ich würde aber fast vermuten, dass es eher so ist, dass diese Art von Betrachtungen aufgesetzt werden auf Ergebnisse, die von der reinen Modellierer-Seite her geliefert werden, und dass dann Kolleginnen und Kollegen zum Beispiel aus der Soziologie damit weitere Forschung betreiben. Bei unserem Exzellenzcluster, das wir in Hamburg haben, dem Clix, da wird ja solche Art von Forschung gemacht. Aber tatsächlich sind dann diese, ich will mal sagen, nachgeschalteten Forschungen und Forschungsfragen nicht in dem Sinne rechenintensiv. Die sind nicht jetzt unterstützt durch mathematische Modellierungen, mindestens nicht durch welche, die auf Hochleistungsrechnern zur Ausführung gebracht werden. Und das ist natürlich eigentlich auch das Spannende und damit auch ein zusätzlicher Rechtfertigungsgrund für die Archivierung dieser großen Datenmengen, dass es einfach Nachnutzungen dieser Daten gibt. Wir sind ja jetzt im Zeitalter der datenintensiven Forschung und dann ist natürlich sehr spannend: Was kann ein Dritter jetzt mit solchen Daten anfangen? Also früher war es ja zum Beispiel so, dass diskutiert wurde: Versicherungen könnten diese Daten vielleicht für irgendwas benutzen. Als ich hier diese Stelle angetreten habe, da war immer so die Idee, dass man vielleicht diese Art der Kommerzialisierung machen könnte von den Daten, hat sich dann hintennach herausgestellt, dass das gar nicht mit unseren Satzungszwecken vereinbar ist. Also da gibt es dann mal an der Stelle noch eine Grenze. Aber zum Beispiel, wenn man nimmt, sagen wir mal, die Tourismusindustrie, wenn irgendeine Region in den Alpen sich die Frage stellt: Was werde ich denn in 20 Jahren meinen Touristen vielleicht an Freizeitbeschäftigungen bieten können, dann sind diese Klimamodelldaten, die erzeugt werden, durchaus eine gute Ausgangsbasis. Und das heißt, man kann darauf aufsetzend jetzt zum Beispiel betriebswirtschaftliche Analysen machen, soziologische Analysen machen, und das ist natürlich sehr spannend.

**[pgg]:** Klingt so, als ob die Nutzercommunity immer größer würde. Wie groß ist sie denn aktuell? Und stimmt das, dass es immer mehr werden, die dann auch irgendwie sich in Hamburg andocken?

**[Ludwig]:** Ja, das mit den Zahlen ist natürlich schwierig. Also ich denke, dass wir an reinen Benutzern auf dem System irgendwas haben, was so an die 1000 mittlerweile rangeht, an einzelnen Projekten, die da gerechnet werden. Da gibt es natürlich dann ganz große Prominente, die jetzt diese neuesten Codes, zum Beispiel diesen Icon-Code, mit dem man auch die Niederschläge berechnen kann und so, die so etwas machen. Tatsächlich, glaube ich, ein Zeichen dafür, dass die Verbindungen enger werden und mehr Wissenschaften auch kooperieren können, ein Zeichen dafür findet man darin, dass wir zum Beispiel jetzt im letzten Jahr unsere Satzung geändert haben, weil, wir haben unsere Satzung dahingehend geändert, dass wir die Daten, die wir haben, an andere Forschungscommunitys auch weitergeben können und auch mit deren Daten operieren können. Ja, weil sprechen wir über Medizin zum Beispiel: Die Veränderung des Klimas hat ja auch beispielsweise gesundheitliche Auswirkungen auf die Menschen, die jetzt in einer bestimmten Region leben. Und man würde in dem Bereich vielleicht gerne Forschung machen, würde da eine Kooperation machen. Nach unserer früheren Satzung, wo drinstand, ja wir dürfen nur Erdsystemforschung unterstützen, könnten wir mit solchen Kollegen überhaupt nicht zusammenarbeiten. Ja, wir dürften von denen keine Daten bei uns lagern, die dürften eigentlich keine Daten von uns verwenden. In dem heutigen wissenschaftlichen Gefüge ist es völlig absurd, weil, es wird einfach so transdisziplinär, interdisziplinär gearbeitet. Und ja, das wird immer mehr. Wir wissen aber zum Teil ja auch gar nicht, wer unsere Daten nachverarbeitet. Die werden über Portale frei zur Verfügung gestellt, unentgeltlich. Und die Leute registrieren sich da, aber wir haben zum Beispiel keine Statistik darüber, aus welchen Wissenschaftsbereichen die kommen. Die allermeisten sind sicher aus dem Bereich Klima- und Erdsystemforschung, aber garantiert sind auch andere.

**[pgg]:** Wie sieht das mit der internationalen Einbindung des Hamburger Standorts aus? Die Klimaforschung selber ist ja international, das Klima ist ja irgendwie auch international oder transnational. Kann man das erklären, wie diese Einbindung dieser Communitys in so was wie eine Weltcommunity aussieht?

**[Ludwig]:** Ja, also dann nehmen wir mal die Community, in der ich mich bewege. Das ist dann so die Community, die die Serviceleistung macht für die Klimaforscher und gleichzeitig aber auch im Bereich Hochleistungsrechnen kompetent ist. Und da gibt es natürlich eine ganze Schiene, also jenseits der nationalen Verflechtungen, die im europäischen Bereich stattfindet. Also im europäischen Bereich ist ja auch sehr stark das ECMWF, das seinen Sitz in England hat, für die Mittelfristwettervorhersage. Metoffice in England, das ECMWF hat seinen Rechner in Bologna stehen, also ist alles sehr europäisch verteilt. In Frankreich gibt es die entsprechenden Institutionen. Also wir, das bedeutet in dem Fall das Max-Planck-Institut für Meteorologie, der Deutsche Wetterdienst, dann das KIT, also Karlsruhe, und das DKZ – wir kooperieren zum Beispiel mit der Schweiz, mit der ETH Zürich und mit dem CSCS in Lugano in Bezug auf die Verbesserung von Codes. Und dann natürlich auch die Erzielung neuer wissenschaftlicher Erkenntnisse. Also das findet natürlich nicht am DKZ statt, sondern beim DWD und beim Max-Planck-Institut. Aber alle zusammen bemühen wir uns um diese Codeverbesserung. Insofern gibt es ein sehr enges europäisches Geflecht. Und darüber hinaus gibt es ein nicht ganz so enges, aber durchaus natürlich auch ein internationales Geflecht. Und jetzt in Kürze ist hier Konferenz. Dann besucht uns der Kollege vom NK in Boulder, diesem National Center for Atmospheric Research. Das ist so von der Konstruktion her wie das DKZ und das Max-Planck-Institut zusammen. Also Forscher und Rechenzentrum in einer Einheit. Und mit denen haben wir engere Kontakte. Da geht es dann um Fragen, wie Daten visualisiert werden, wie Programme

beschleunigt werden, wie man den Rechner überhaupt kauft und aufbaut. Ja, das sind also auch diese ganz pragmatischen Fragen. Und das DKZ zum Beispiel, für sich genommen, ohne die Nutzer, hat darüber hinaus als Kontaktpartner alle Hersteller von Hochleistungsrechnern. Wenn wir einen neuen Rechner kaufen, machen wir entsprechende Besichtigungs- und Erkundungstouren in den Vereinigten Staaten, in Japan und in China. Wir sprechen mit allen Herstellern von Hochleistungsrechnern über Produktreihen, um dann zu sehen, was für unsere Kundschaft günstig wäre.

**[pgg]:** Kann man sich das so vorstellen, dass die Bekämpfung des Klimawandels tatsächlich alle internationalen Akteure friedlich und kollaborativ zusammenbringt? Oder gibt es da auch Konkurrenzen?

**[Ludwig]:** Ach, ich denke, es ist immer so: Konkurrenz ist wichtig. Konkurrenz regt an, belebt das Geschäft und führt dazu, dass die Leute sich nicht auf ihren Erfolgen ausruhen. Wenn man zum Beispiel guckt, was im IPCC-Bericht gemacht wird, dann ist ja Band 1 von diesen drei Bänden, der ist ja dem Modellvergleich gewidmet. Dieses, was sich CMIP nennt, Coupled Model Intercomparison Project, also ein Projekt zum Vergleich gekoppelter Modelle. Gekoppelt bedeutet immer Ozean und Atmosphäre. Und da werden ja über die Welt hinweg über 20 Institutionen mit ihren Modellen jeweils teilnehmen. Und es wird verglichen, ob die Aussagekraft der Modelle bei vorgegebenen Randbedingungen, also Staatsszenarien und Laufzeiten und so was, ob die Aussagekräfte dieser Modelle vergleichbar sind und ob es Modelle gibt, die da rausfallen. Also viele Modelle sind weitestgehend unabhängig voneinander entwickelt worden. Aber die Forscherinnen und Forscher sind natürlich auch mobil. Ja, es gibt genügend, die mal nach Amerika gehen oder solche, die rüberkommen. Dann wandern auch die Ideen mit oder es wandern sogar auch Softwareteile mit. Also so ganz hundertprozentig unabhängig sind sie nicht entwickelt worden. Aber in dem Zusammenhang ist es schon spannend, dann auch zu sehen: Welche Aussagekraft haben die einzelnen Modelle, die an Standorten entwickelt werden? Und vielleicht auch die Frage: Kommen sie schneller oder energieärmer zu ihrem Ergebnis? Das ist heutzutage ja auch ein Aspekt. Und insofern: NK veranstaltet jetzt im September wieder seinen Workshop, da werden weltweit Leute teilnehmen, die genau sowas machen. Insofern ist das eben Kooperation und Konkurrenz in einem. Und das ist eigentlich auch sehr schön.

**[mg]:** Das Stichwort wie energieeffizient kommt mal zum Ergebnis. Das ist ja eigentlich auch nochmal ein interessanter Punkt. Wie groß ist denn da die Spanne?

**[Ludwig]:** Ja, welchen Stromverbrauch die haben, um zu einem bestimmten Ergebnis zu kommen. Das wurde jetzt noch nicht unabhängig analysiert. Was wir hier bei uns machen, ist, dass wir natürlich versuchen, für ein einzelnes gegebenes Modell, das bei uns läuft, den Stromverbrauch möglichst gering zu halten. Also dadurch, dass wir Tuning machen, des Modellcodes, dadurch, dass wir unsere Infrastruktur passend bauen. Ich selber habe in diesem Bereich Energieeffizienz geforscht. Und da waren Anfang der Zehnerjahre auch einige Projekte. Gab es zum Beispiel durchaus mal ein Projekt zu der Frage: Welche mathematische Umsetzung einer konkreten Fragestellung ist eigentlich die energieeffizientere? Man braucht eine bestimmte Mathematik, um zum Ergebnis zu kommen. Dann hat man drei verschiedene Varianten, wie man diese Mathematik in einen Programmcode gießen kann. Dann lässt man die laufen und misst sie aus. Und dann sind die natürlich unterschiedlich in ihrem Energieverbrauch. Und da kann man durchaus Rückschlüsse darauf ziehen, was



man vielleicht tun und lassen sollte. Es ist jetzt noch nicht so weit, dass man tatsächlich der Mathematik das einfach so zuschreiben kann und kann sagen: Das ist günstig, das ist ungünstig. Aber Klimaforschung, zum Beispiel Klimacodes als solche, laufen auf Rechnern nicht besonders energieeffizient. Leider. Da gibt es Codes, zum Beispiel aus dem Bereich der Bioinformatik, da werden sehr viele Rechenoperationen auf kleinen Datenmengen gemacht. Sowas ist energetisch in einem Rechner günstiger. Und beim Klima sind die Datenmengen zwar auch groß, aber die Mathematik, die pro Datenpunkt angewendet wird, ist eher überschaubar. Bei uns sind die Rechner in der Folge im Großteil ihrer Zeit damit beschäftigt, Daten aus dem Hauptspeicher zum Prozessor zu schaufeln und wieder zurück. Und in der Zeit, wo sie solche Aktivitäten machen, können sie nicht rechnen, brauchen aber trotzdem Strom. Das heißt sozusagen: Rechenoperation pro Kilowatt ist beim Klima, wenn man Klima rechnet, ungünstiger als wenn man zum Beispiel in der Pharmazie jetzt so Proteinfaltungen ausrechnet. Tatsächlich gibt es da communityspezifische Unterschiede dann.

**[mg]:** Wie ist denn so die Energiebilanz insgesamt, würden Sie sagen, die ist noch so schlecht, dass es da dringenden Optimierungsbedarf gibt?

**[Ludwig]:** Also als Informatiker rechnen wir ja immer in folgender Weise. Wir gucken: So ein Prozessor, der hat innen drin diese sogenannten Rechenwerke, mit denen er die mathematischen Operationen macht. Und wir hätten natürlich am liebsten, dass der Prozessor 100 Prozent seiner Zeit nur Mathematik macht, weil, wir kaufen ihn ja dafür. Also, das wäre irgendwie super. Das wäre, wie wenn man sagt, man kauft sich einen Porsche. Quasi ist es bei uns so: Wir kaufen einen Porsche und fahren ihn ständig in der Stadt, weil die Ausbeute, die wir mit unseren Klimacodes haben, die liegt etwa im Bereich 5 Prozent. Also 5 Prozent der Operationen, die so ein Rechenchip pro Sekunde machen könnte, gehen tatsächlich in den Klimacode rein. Und da sind halt die genannten Codes, sagen wir mal Proteinfaltung von der Bioinformatik, die kommen dann da mal so in den Bereich 20 Prozent, dann ist es deutlich besser. 100 Prozent erreicht man dadurch nie, weil, neben der Mathematik hat so ein Code ja auch Sprünge im Code oder ich muss irgendwelche Daten aus dem Hauptspeicher holen. Also, es ist klar, dass ich nicht 100 Prozent der Zeit rechnen kann, aber beim Klima ist relativ wenig Rechnung. Und wir versuchen, durch Verbesserung der Codes diesen Prozentfaktor zu erhöhen. Der Nutzen, der dadurch entsteht, ist, dass wir für in etwa denselben elektrischen Stromverbrauch eine kürzere Laufzeit des Programms dann haben und dadurch sozusagen das Programm energiesparender wird. Also ich meine, das ist der ökonomische und ökologische Gewinn. Und der Gewinn für den Benutzer oder die Benutzerin ist natürlich, dass das Ergebnis früher da ist.

**[pgg]:** Sind auch die Rechner selber, also die Rechnerarchitekturen und ihre Bauweise energiemäßig unterschiedlich angelegt? Hat sich da was getan in den letzten Jahren?

**[Ludwig]:** Ja, da hat sich sehr viel getan. Also, ich bin ja 2009 nach Hamburg gekommen, da hatte ich aus Heidelberg, wo ich vorher war, so ein Projekt zum Thema Energieeffizienz mitgenommen. Und als ich hier angekommen bin, hatten die gerade einen IBM Rechner in Betrieb genommen, im Jahr 2009. Und die Entwicklung war ja dann aus dem Jahr 2007 etwa. Die Modelle zu dieser Zeit, die hatten überhaupt keine Mechanismen für Energieeffizienz. Damals war das einfach so, dass man gesagt hat: Ja, naja, Hochleistungsrechnen, das ist so die Formel eins des Rechnens und es braucht halt mehr. Und so ist es. Aber es war dann absehbar, so ab den 2000ern: Mit den Stromkosten kann so eine Haltung natürlich nicht einfach fortgesetzt werden. Also, wir

haben dann den nächsten Rechner beschafft. Das ist der, den wir jetzt ausschalten. Den haben wir beschafft im Jahr 2015. Und in diesem Rechner waren dann schon verbaut die ganzen Mechanismen, die wir auch in Notebooks haben, weil Notebooks als potenziell batteriebetriebene Geräte andere Konstrukte im Prozessor haben, insbesondere so Konstrukte, die zum Beispiel, wenn die Rechenlast auf dem Rechner geringer ist, dann auf dem Prozessor die Spannung runterschalten und die Frequenz runterschalten, was zu einer deutlichen Energieverbrauchsreduktion führt. Und diese ganzen Mechanismen, das gab es bei dieser IBM Maschine überhaupt nicht. Die hat immer 100 Prozent Last erzeugt, also am Stromnetz, egal, ob da jetzt Programme drauf liefen oder nicht. Das ist bei den neueren Rechnern absolut nicht mehr der Fall. Also das ist mal das eine, diese ganzen Energiesparmechanismen, die in batteriebetriebenen Geräten Standard waren. Die hat man einfach in die Prozessoren fürs Hochleistungsrechnen auch eingebaut. Und dann hat man darüber hinaus aber so Infrastrukturmaßnahmen gemacht, die wir auch nutzen. Zum Beispiel haben wir sowas, das nennt sich Hochtemperaturflüssigkeitskühlung. Da können wir die Rechner durch durchströmende Flüssigkeit kühlen. Früher hat man Rechner immer nur luftgekühlt, aber Luftkühlung ist sehr, sehr ineffizient. Und bei dieser Flüssigkeitskühlung, bei dieser Hochtemperaturvariante sogar, da ist es uns erlaubt, dass wir mit bis zu 40 Grad Kühlflüssigkeit in diese Rechner hineingehen. Und dann kommt die Flüssigkeit vielleicht mit bis zu 50 Grad raus und wir müssen sie halt auf 40 Grad runterbringen. Und in Hamburg, hier bei uns am Dach, hatte es noch nie mehr als 38,5. Das heißt, wir können dieses Absenken der Temperatur passiv erreichen. Wir pumpen einfach die heiße Flüssigkeit aufs Dach und dann macht die umgebende Luft ihre Arbeit. Das sind so langsam drehende Ventilatorenblätter oben. Und damit erreichen wir, dass wir unterhalb dieser 40 Grad kommen, die vorgeschrieben sind. Und dann läuft die Flüssigkeit wieder in den Rechner rein. Auf diese Methode spart man sich die ansonsten sehr energiehungrigen Kühlanlagen. So eine Kühlanlage, um diese IT zu kühlen, die kann leicht nochmal 30 bis 50 Prozent des Stromverbrauchs vom Rechner dazufügen, sodass man dann 1,3-, 1,5-Mal so hohe Stromkosten hat. Und bei dieser Hochtemperaturflüssigkeitskühlung, da ist der Zusatzaufwand für die Kühlung, der liegt so im Bereich 4 Prozent vielleicht. Da ist man jetzt quasi fast schon an einem Optimum. Ich glaube, die nächste optimale Stufe wäre, man stellt es irgendwo am Nordpol einfach ins Freie und es ist kalt. Dann braucht man gar nichts mehr kühlen, da braucht man auch keine Flüssigkeit durchpumpen.

**[pgg]:** Ich habe nochmal eine ganz andere Frage. Es ist sehr eindrucksvoll, wie eng die verschiedenen Leistungsbereiche verknüpft sind, die nötig sind, um wirklich so komplexe wissenschaftliche Ergebnisse zu produzieren. Wird das dann am Ende alles auch so mit abgebildet? Also werden die Leute auch mitgenannt, die fürs Rechnen und die Maschinen verantwortlich sind, wenn nachher die wissenschaftliche Leistung publiziert wird? Oder sind das doch nur die Nutzerinnen und Nutzer des DKRZ, die dann nachher auf den Aufsätzen draufstehen?

**[Ludwig]:** Ja, es ist schon eher Letzteres. Und tatsächlich ist das immer ein Problem, wenn es darum geht, beim Ministerium vorzusprechen für neue Gelder für die Infrastruktur. Also wir bemühen uns immer, eine Lösung auf die Frage zu finden: Wie können wir herausfinden, was die Wissenschaftlerinnen und Wissenschaftler dank unserer Infrastruktur an wissenschaftlicher Wertschöpfung erzielt haben? Und da machen wir manchmal so bibliometrische Analysen, wo wir dann gucken: Wie häufig wird das DKZ genannt? Nutzer auf unseren Maschinen, die Rechenzeit aufgrund des wissenschaftlichen Lenkungsausschusses bekommen haben, die müssen so ein

Acknowledgement in ihre Paper reinschreiben. Und dann kann man praktisch durch so Analysen herausfinden: Ja, wie viele Veröffentlichungen hat es denn da gegeben? Das ist für uns durchaus ein wichtiger Punkt. Wir stehen jetzt exakt, jetzt im Moment, wieder an dieser Frage, weil wir einen neuen Antrag schreiben müssen für die nächste Generation von Rechner, die wir beschaffen. Und da muss natürlich irgendwie so eine Veröffentlichungsliste mit rein. Und wir können bei allen anklopfen und sagen: Gebt uns mal eure Veröffentlichungen. Das tun wir auch, ja, aber das funktioniert auch nicht so hundertprozentig. Besser ist immer, wenn wir zusätzlich noch Analysen machen. Und tatsächlich ist es so: Nein, wir werden da nicht systematisch genannt. Und das ist durchaus ein Problem. Das betrifft aber alle Forschungsinfrastrukturen, weil diese sehr teuren Forschungsinfrastrukturen dadurch nicht so ganz die Sichtbarkeit haben, welchen ermöglichenden Faktor sie eigentlich spielen bei dann der wissenschaftlichen Wertschöpfung.

**[mg]:** Das dachte ich auch gerade, ob es nicht vielleicht wichtig wäre, dass diese Infrastrukturen mehr Öffentlichkeit haben. Also es ist ja auffällig auch, dass das Klimarechenzentrum diese zentrale Bedeutung hat für die Klimaforschung. Und der Klimawandel ja nun auch relativ viel Aufmerksamkeit bekommt, auch wenn vielleicht die Maßnahmen immer noch zu wünschen übriglassen. Also das scheint ja sozusagen, dass das nicht nur, sag ich mal, in den Medien nicht genug wahrgenommen wird, sondern auch in der Politik. Habe ich das richtig verstanden oder ist es nicht so dramatisch?

**[Ludwig]:** Also ich glaube, ein Stück weit ist es tatsächlich einfach so, wie Sie sagen. Und das ist jetzt nicht beschränkt aufs DKZ, sondern es ist in dem Land insgesamt halt so, dass die Bedeutung von IT und die Bedeutung des Digitalen, wenn man mal das so sagen, nicht genügend berücksichtigt wird. Das findet sich in der schulischen Ausbildung, das findet sich bei Behörden und so weiter. Also, da wird natürlich jetzt viel darüber geredet, was wir alles aufholen müssen. Aber allein was wir alles aufholen müssen, zeigt schon, dass die Berücksichtigung der Wichtigkeit dieser Konzepte, das hat sowohl in der Bevölkerung als auch in der Politik gar nicht stattgefunden. Es wird auch schön dokumentiert durch einen Satz, mit dem die Leute gerne kokettieren, indem sie sagen: Ich war immer schlecht in Mathe. Schlecht in Mathe zu sein, ist in Deutschland irgendwie schick. Und sozusagen ist das natürlich ungünstig für eine moderne Gesellschaft, bei der so viel auf IT dann basiert und so viel auf Digitalisierung basiert. Dieses Nicht-Wahrnehmen, dass diese Infrastruktur halt sehr wichtig ist.

**[mg]:** Da kann ich an der Stelle vielleicht gleich mal den Hörern noch den Tipp geben, tatsächlich mal auf Ihre Webseite zu schauen, weil, die ist wahnsinnig informativ und Sie bieten ja auch Führungen an. Wie stark wird das denn nachgefragt?

**[Ludwig]:** Das wird sehr stark nachgefragt. Also, wir hatten ja jetzt dann zwei Jahre, in denen sehr wenig passiert ist. Jetzt hier, also tatsächlich in diesen Wochen, fängt es hier wieder an zu boomen. Das letzte normale Jahr war das Jahr 2019. Im Jahr 2019 haben wir etwa 1500 Gäste am DKZ gehabt. Der größte Teil davon, würde ich sagen, sind die Schulklassen. Aber dann sind es natürlich auch Gruppen von Universitäten. Dann sind es zum Beispiel Rotaria. Wir hatten hier schon den Deutschen Hausfrauenbund. Gelegentlich wird auch mal irgendwie eine Filmszene bei uns gedreht. Wir hatten schon den Tatort im Haus. Die kriegen natürlich keine Fortbildung über Klimaergebnisse. Aber alle anderen kommen, weil sie meistens so eine Kombination erfahren möchten, über das, was jetzt eigentlich die Ergebnisse der

modernen Klimamodellierung sind. Wir haben im Haus natürlich viele Leute, die auch Geowissenschaften studiert haben, gerade die in der Öffentlichkeitsarbeit, die dann sowas kommunizieren können. Da bin ich als Informatiker immer nicht so richtig gut dazu geeignet. Und wir teilen uns dann den Job immer dahingehend. Der Herr Böttinger, der die Öffentlichkeitsarbeit leitet, macht dann den Teil übers Klima. Und ich mache dann die Rechnerraumführung und erzähle den Leuten den Teil über die IT und die Informatik. Und manchmal ist klar, wir hatten hier von der Bundeswehrhochschule Informatik-Studierende, die haben sich mehr für den Rechner und die Details interessiert. Wir gestalten das dann je nach Kundenkreis, wenn ich es mal so sagen darf, flexibel. Und dann bekommen die immer die Informationen, die sie gerne hätten. Und ja, wir haben wirklich viele Besucher hier.

**[mg]:** Ich habe noch eine neugierige Frage, die ein bisschen auch auf zukünftige Anschaffungen anspielt. Denken Sie oder können Sie schon sagen, ob das DKRZ seinen eigenen Quantencomputer haben wird eines Tages? Oder planen Sie da Kooperationen mit anderen Partnern? Spielt das überhaupt eine Rolle?

**[Ludwig]:** Aktuell beobachten wir so ein bisschen, was sich tut. Aber wir brauchen uns nicht so sehr einzumischen, weil, es tut sich zu wenig noch. Zu wenig, was bei uns nutzbar wäre. Ich habe ja am Anfang erwähnt, dass wir an und für sich immer relativ konservativ die Hardware-Architekturen beschaffen, weil wir diese großen Codes, die haben ja manchmal eine Million Zeilen Code und mehr, nicht einfach umschreiben können. Beim Quantencomputer ist es Stand heute so, dass die Frage, wie man ihn programmiert, noch ungeklärt ist. Insbesondere auch so die Frage, wie man jetzt ein partielles Differenzialgleichungssystem, und darauf beruhen die meisten unserer Anwendungen, wie man so ein mathematisches Konstrukt auf den Quantencomputer bringt. Daran wird geforscht und da gibt es auch Fortschritte, aber keine, die wir jetzt direkt verwenden könnten. Ich bin kein Experte dafür, aber ich denke, dass beim Quantencomputer die Situation schon so sein könnte, dass man vielleicht durch eine Entdeckung, die jemand macht, doch überraschend schnell zu einer Nutzung kommen kann und nicht notwendigerweise nur die Entwicklung abwarten muss. Also ich glaube, wenn man die Entwicklung abwartet, dauert es sicher noch 20, 30 Jahre, bis das bei uns eine Rolle spielt, jenseits meines Zeithorizonts hier. Aber bei dieser komplexen Materie könnte es sein, dass durch einzelne Entdeckungen, die jemand macht, sich vielleicht diese Zeitspanne verkürzt. Wir gehen momentan davon aus, dass so ein Quantencomputer, genau wie diese schon genannten Grafikkarten, so eine Art Beschleunigerhardware sein könnte für einzelne Abschnitte in den Berechnungen, die wir dann auf den Quantencomputer auslagern würden, und der Hauptteil läuft auf einem klassischen Parallelrechner. Aber selbst so einen kombinierten Einsatz, glaube ich, den wird man nicht innerhalb der nächsten zehn Jahre schon direkt sehen. Das wird noch etwas dauern.

**[mg]:** Was verspricht man sich denn, ich komme jetzt einfach mal mit so Schlagworten, was verspricht man sich denn vom maschinellen Lernen? Sie hatten ja schon gesagt, das ist vor allem für die Wetterprognosen interessant, aber Sie bieten ja maschinelles Lernen auch als Service an. Welche Rolle wird das denn spielen?

**[Ludwig]:** Das wird eine interessante Sache, weil da habe ich eine Meinung als DKZ-Chef und eine Meinung als Wissenschaftler. Also als DKZ-Chef, glaube ich, ist es so: Wir werden das einfach an sehr vielen Stellen einbauen und einbauen können. Und wir haben hier mittlerweile zwei Gruppen. Die einen, die machen genau das, was Sie

sagten, die sind in einem serviceorientierten Umfeld der Helmholtz-Gemeinschaft ansässig. Und die andere Gruppe macht mehr im Bereich Forschung etwas. Und zwar Forschung jetzt, wie man maschinelles Lernen im Bereich Klimamodellierung einbringt. Was wir hier nicht machen, ist jetzt Methodenforschung zum maschinellen Lernen, weil das überlässt man eigentlich den Informatikern und den Spezialisten, die es hierfür gibt. Also wir versuchen das dann anzuwenden. Und meine persönliche Meinung ist so als Wissenschaftler, dass ich denke: Wir haben in all diesen Fällen eigentlich ein Problem mit der Erklärbarkeit. Ja, weil das klassische maschinelle Lernen, das wir heute finden, dieses Deep Learning, das ist ja im Wesentlichen eigentlich eine Art von statistischem Verfahren. Das lernt bestimmte Sachverhalte und kann sie dann auf andere anwenden. Und man darf nicht dann hinten nach die Frage stellen: Warum ist das jetzt genau deine Antwort? Weil dann muss das Ding natürlich sagen: Ja, die Gewichte in meinen künstlichen Neuronen, die waren halt so und so. Und das ist jetzt was, was für einen Forscher, der daran interessiert ist 'warum ändert sich das Klima in dieser oder jener Weise?', natürlich unbefriedigend ist. Weil, dem ist es natürlich lieber: Er kann ein mathematisches Gleichungssystem hinschreiben und kann sagen, weil Druck und Temperatur und diese Sachen in dem und dem Verhältnis zueinanderstehen und Naturgesetze gelten, deshalb verhält sich das mit dem Wetter so und so. Und jetzt schreibe ich euch das hin und dann machen wir ein Programm draus und dann rechnen wir das aus. Würde jetzt ein maschinelles Lernverfahren das wesentlich schneller können und es wäre aber nicht erklärbar, warum das Ergebnis valide ist, wäre das aus wissenschaftlicher Sicht sehr unbefriedigend. Und das unterscheidet Wetter und Klima, weil jetzt zum Beispiel bei der Wettervorhersage, ich mache immer den Witz und sage: Die ist ja eh falsch, und wenn sie durch maschinelles Lernen ein bisschen falscher wird, ist es eigentlich egal. Da ist eher die Frage: Wie schnell habe ich sie? Weil da kann ich nämlich vielleicht mit maschinellem Lernen eine eh falsche Wettervorhersage in einem Drittel der Zeit berechnen. Das heißt letztendlich auch mit einem Drittel von der elektrischen Leistung. Das ist natürlich total interessant. Brauche ich viel weniger Strom oder brauche ich einen kleineren Rechner und habe vielleicht nur einen kleinen Qualitätsverlust. Aber wenn ich maschinelles Lernen ins Klima, in die Klimamodellierung einbringe, habe ich halt diesen Verlust an Erklärbarkeit. Und den wird man nicht an allen Stellen akzeptieren wollen. Da gibt es so bestimmte Bereiche, wo man das jetzt heutzutage integriert und dann guckt man mal, was geht. Aber wir sind tatsächlich mit maschinellem Lernen und der Integration in die Klimamodellierung nicht so weit wie bei der Integration in zum Beispiel Produktionsbetrieb bei Wettervorhersage. Wenn ich verstehen will, wie das Wetter funktioniert, ist mir eigentlich maschinelles Lernen eher im Weg – wissenschaftstheoretisch extrem spannend.

**[mg]:** Der Blick nochmal in die Politik. Also, es ist ja schon so, dass die Klimaforschung vor allem in die Zukunft schaut, versucht, die zukünftigen Entwicklungen zu modellieren, vielleicht auch den Klimawandel vorherzusagen, Szenarien zu berechnen. Der IPCC-Report versteht sich ja auch als Grundlage für politische Entscheidungen. Und die öffentliche Hand leistet sich im Grunde diese Wissenschaft, also stellt diese Ressourcen zur Verfügung, auch so spezialisiert, wie das jetzt für das Deutsche Klimarechenzentrum der Fall ist. Fragen Sie sich manchmal, warum das so ist, wenn dann die Ergebnisse so wenig in die Umsetzung gehen, in die politische? Oder ist das so eine öffentliche Einschätzung, die eigentlich nichtzutreffend ist?

**[Ludwig]:** Da hat sich viel geändert. Ja, ich glaube, dass die Ergebnisse schon sehr wohl zu gesellschaftlichem Handeln führen. Und dann würde ich das vielleicht auch lieber

kontrastieren wollen zu anderen Wissenschaftsdomänen, von denen ich weiß, dass sie zum Beispiel enorm viel Rechenzeit verbrauchen. Nehmen wir mal an, also ich mag das eigentlich, aber nehmen wir mal die Astronomen mit Galaxie-Kollisionen. Die rechnen Galaxie-Kollisionen auf irgendwelchen Supercomputern in Jülich, verbrauchen Unmengen an Strom dafür. Und da muss man auch sagen: Okay, wenn wir in vier Milliarden Jahren mit Andromeda kollidieren, schön zu wissen, aber so what? Und gemessen daran, glaube ich, ist die Bilanz zwischen Investition bei jetzt der Klimaforschung und gesellschaftlichem Nutzen sehr, sehr gut.

**[pgg]:** Ja, gerade auch diese interessante Perspektive, dass so kleinräumig interessierte Kunden was davon haben von guten Szenarien und dass also nicht nur verantwortliche Kommunen, die sich um Veränderungen in ihrem direkten Verantwortungsbereich irgendwie kümmern müssen, sondern eben auch Unternehmen oder zivilgesellschaftliche Gruppen da im Detail informieren können, das, glaube ich, sichert diesen Daten eine enorme Zukunft.

**[Ludwig]:** Das ist so, ja. Ich kann mal ein Beispiel nennen, eigentlich sind es sogar zwei Beispiele. Die werden natürlich leider komplett unterm Deckel gehalten. Und zwar sind es die letzte Installation vom Metoffice in England und die anstehende. Die letzte Installation hatte einen Wert von 100 Millionen Euro. Und die jetzt anstehende, da gab es letztes Jahr, vorletztes Jahr irgendwie die Ausschreibung, hat Microsoft gewonnen. Da geht es um zwei Rechnergenerationen, die da betrieben werden. Und der Vertrag hat, das ging damals durch die Presse, einen Wert von 1,4 Milliarden Euro. 1,4 Milliarden aggregierte Summen. Wahrscheinlich ist das inklusive Strom, inklusive Rechner, inklusive Gebäude und allem. So, und dann fragt man sich: Warum so viel Geld? Und bei beiden Ausschreibungen wurde Folgendes gemacht, aber das ist genau der Punkt, an dem man nicht an die Details kommt: Es gab im Vorfeld Analysen darüber, wie man durch zum Beispiel Vorhersagen von Extremwetter und Klimaveränderungen volkswirtschaftliche Schäden vermeiden kann, wenn man einen bestimmten Rechner einer bestimmten Leistungsklasse hat. Und gerüchteweise hört man, dass für diesen Rechner, für den es jetzt diese 1,4 Milliarden hinlegen, diese Studie irgendwie gezeigt hat, man könnte volkswirtschaftliche Schäden in Höhe von 13 Milliarden Euro vermeiden. Wenn man so eine Rechnerinfrastruktur kauft, betreibt, auswertet und handelt. Ob das dann so gemacht wird? Ich weiß es nicht. Ich weiß auch nicht, was in dieser Studie drinsteht, aber es ist jetzt schon die zweite Generation, der bei der Beschaffung und bei der Zurverfügungstellung dieser Geldsummen so eine Analyse zugrunde liegt. Und ich habe so Anfang der 10er Jahre mehrfach Workshops selber gemacht zum Thema Kosten-Nutzen-Analysen von Forschungsinfrastrukturen. Denn es ist ja schon die Frage: Der Steuerzahler zahlt jetzt das hier, und was bringt es jetzt dem Steuerzahler zum Beispiel? Und es gibt durchaus Analysen dann, ich habe es leider nicht auf meinem Tisch, aber ich fände es sehr spannend. Es gibt durchaus Analysen, die halt zeigen, dass man viel mehr damit sparen kann, als es tatsächlich an Kosten verursacht. Und dass deshalb diese Investitionen rechtfertigbar sind zum Nutzen der Gesellschaft.

**[mg]:** Das heißt, der Frust in der Klimaforschung ist gar nicht so groß, wie man manchmal denken könnte so als Außenstehender?

**[Ludwig]:** Also, das müssten Sie meine Kolleginnen und Kollegen fragen. Ich glaube, es ist immer so eine Kombination, denn ich denke, jetzt gerade bei denen, die so kundig sind, also Leute, die diese Dinge erforscht haben und sie in die Öffentlichkeit auch

tragen, macht ja auch nicht jeder Forscher, weil manche Forscherinnen und Forscher sind einfach nur daran interessiert, diesen Gegenstand zu erforschen und fertig. Aber die, die das in die Gesellschaft tragen, die machen das mit Enthusiasmus und mit Besorgnis. Und ich glaube, sie sind auch immer wieder frustriert. Aber es wird trotzdem auch zunehmend mehr gehandelt, also agiert. Immer noch zu langsam, aber halt trotzdem im Aufschwung. Ja, und es sind alles Effekte dieses ständigen, gebetsmühlenmäßigen Warnens, das die Forscherinnen und Forscher machen. Und ja, und ich denke, es ist eine Mischung aus Frustration manchmal und trotzdem Freude auch, was beizutragen.

*[Der Abspann mit Musik beginnt.]*

**[mg]:** Und damit ist dieses Digitalgespräch zu Ende. Vielen Dank an Thomas Ludwig vom Deutschen Klimarechenzentrum für das spannende Gespräch und die vielen interessanten Eindrücke und Einblicke. Viele Grüße nach Hamburg. Und Dank auch Ihnen, liebe Zuhörerinnen und Zuhörer, dass Sie uns diesmal wieder Ihre Aufmerksamkeit geschenkt haben. Wenn Sie mögen, hören wir uns in drei Wochen wieder zur nächsten Folge des Digitalgesprächs, dem Podcast von ZEVEDI, dem Zentrum verantwortungsbewusste Digitalisierung.



This work is licensed under CC BY-NC-ND 4.0. To view a copy of this license, visit <https://creativecommons.org/licenses/by-nc-nd/4.0/>