



AKPIK-DIGIT

Deutsche Physikalische Gesellschaft: Arbeitskreis Physik, moderne Informationstechnologie und Künstliche Intelligenz akpik.dpg-physik.de

| | |
|---|----|
| In dieser Ausgabe von AKPIK-DIGIT | S. |
| Herzliche Begrüßung! | 1 |
| Wer ist Mitglied im AKPIK | 2 |
| AKPIK auf Tagungen | 3 |
| Info: Industriegespräche Digitalisierung | 4 |
| Info: Plattform Lernende Systeme | 4 |
| Info: Erforschung von Universum & Materie | 5 |
| Info: Nationale Forschungsdaten-Infrastruktur | 5 |

Längst ist das Exabyte (1.000.000 Terabyte) am CERN in Genf überschritten: **Big Data** ist schon lange in der Physik angekommen.

Warum ist moderne Digitalisierung drängendes Thema der Physik?

Moderne Digitalisierung umfasst vier Bereiche, die in der Forschung arbeitende und industrienahe arbeitende Physikerinnen und Physiker betreffen:

1. Wie holen wir entscheidende Informationen effizient aus Messdaten für neue Erkenntnisse in der Physik?

Algorithmen sind die Sprache, mit der wissenschaftliche Fragen an die Daten gestellt werden: Aktuell liegt ein Fokus auf Entwicklungen in den Bereichen Verbesserungen im parallelen Rechnen und Varianten des maschinellen Lernens (Deep Learning, Informationsfeldtheorie).

2. Wie bewältigen wir demnächst Datenströme von Exabyte pro Jahr?

Dafür sind neue Konzepte für die Zusammenstellung geeigneter Computer-Hardware erforderlich. Außerdem braucht es Steuerprogramme (Workflows), mit denen Daten effizient durch die Computeranlagen geleitet werden, so dass die jeweils optimale Hardware (CPU, GPU, Memory-Bedarf) für die Algorithmen verwendet wird.

Liebe Mitglieder des AKPIK von Martin Erdmann

Im Namen des Vorstands begrüßen wir Sie ganz herzlich im Arbeitskreis AKPIK! Im Sinne unseres Mottos *Bekanntmachen – Verbinden – Weiterbilden* stellen wir Ihnen heute den ersten halbjährlich erscheinenden AKPIK-DIGIT vor.

Digitalisierung ist unter Physikerinnen und Physikern wahrlich kein Fremdwort.

Seit Jahrhunderten werden in unseren Experimenten Messwerte genommen und so verarbeitet, dass wir neue Erkenntnisse daraus ziehen können. Mit immer moderneren Experimenten kommen immer neue Größenordnungen von Datenmengen, die verarbeitet und interpretiert werden müssen.

3. Wie machen wir Forschungsdaten langfristig nutzbar, so dass sie bei neu aufkommenden Fragestellungen weiterverwendbar sind?

In den meisten Fällen werden Messdaten eines Experiments bislang so aufbereitet, dass effiziente Datenanalysen und Publikationen für das Experiment ermöglicht werden. Die neue Initiative für eine nationale Dateninfrastruktur zielt darauf, durch Kombinationen von Messdaten verschiedener Experimente den Erkenntnisfortschritt zu beschleunigen.

4. Wie integrieren wir die modernen Formen der Digitalisierung in die Physik-Ausbildung?

Eine Grundausbildung in Statistik erhalten alle Studierenden für die Datenanalysen der Laborpraktika. Allerdings hat die Bedeutung von Berechnungen und Simulationen am Computer in allen Bereichen der Physik sehr stark zugenommen. Eine verbesserte Ausbildung in der Entwicklung gut strukturierter Quellcodes und effizienter Algorithmen, in Simulationen und Maschinellem Lernen wird derzeit diskutiert und in mehreren Fachbereichen bereits umgesetzt.

Es gibt reichlich nationale und internationale Zusammenschlüsse zu diesen Themen unter

verschiedenen Akronymen (PLS, ErUM, NFDI): Wir stellen Ihnen weiter unten eine Auswahl vor. Ebenso informieren wir Sie über relevante Events (Workshop, Tagung, Diskussionen).

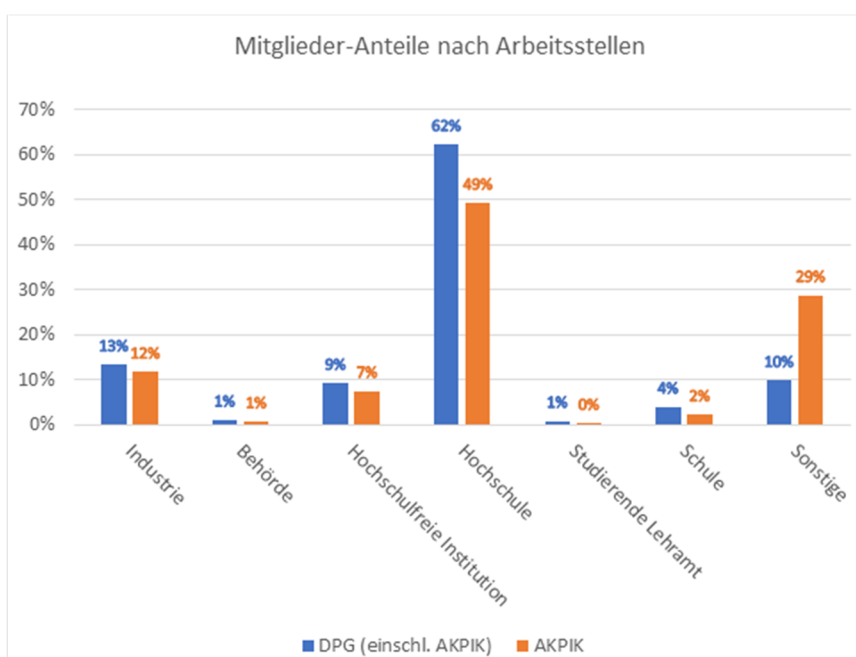
Wer ist Mitglied im AKPIK

von Andreas Hamm und Andreas Fehlner

Der AKPIK hat weniger als zwei Jahre nach seiner Gründung bereits 915 Mitglieder (Stand vom 06.09.2019). Das entspricht etwa 1.6% aller aktiven DPG Mitglieder.

Ein Blick auf die Zusammensetzung des AKPIK nach Arbeitsstellen gibt darüber Auskunft, welche Berufsgruppen besonders stark im AKPIK vertreten sind – absolut betrachtet, aber auch im Vergleich zur Arbeitsstellenverteilung aller DPG-Mitglieder. Abbildung 1 zeigt, welche Anteile auf die verschiedenen Berufsgruppen entfallen – einerseits in der Gesamt-DPG (blaue Säulen), andererseits im AKPIK (orange Säulen).

Grundsätzlich kann man sagen, dass die Arbeitsstellenstruktur im AKPIK nicht wesentlich von der in der Gesamt-DPG abweicht – mit einer bemerkenswerten Ausnahme: Während auch im AKPIK die



Arbeitsstelle „Hochschule“ den mit Abstand größten Anteil ausmacht, ist im AKPIK die Gruppe „Sonstige“ viel stärker als in der Gesamt-DPG. Diese Sammelgruppe umfasst unter anderem Abiturientinnen und Abiturienten, die neu in die DPG eintreten.

Hier deutet sich bereits an, was die Abbildung 2 belegt: Der AKPIK ist ganz besonders bei jungen DPG-Mitgliedern beliebt. Die Altersverteilung im AKPIK (orange Säulen) ist deutlich verschieden von der Altersverteilung in der Gesamt-DPG (blaue Säulen).

Während in der Gesamt-DPG die Gruppe der 31- bis 40-jährigen die stärkste Altersgruppe ist, sind es im AKPIK die Unter-20-jährigen. Die älteren Altersgruppen sind im AKPIK eher schwach vertreten.

Die drei größten Untergruppen im AKPIK sind also:

- StudienanfängerInnen
- ForscherInnen an Hochschulen und hochschulfreien Institutionen, überwiegend am Beginn ihrer Laufbahn
- PhysikerInnen in der Industrie, die meist vor nicht langer Zeit die Hochschule verlassen haben

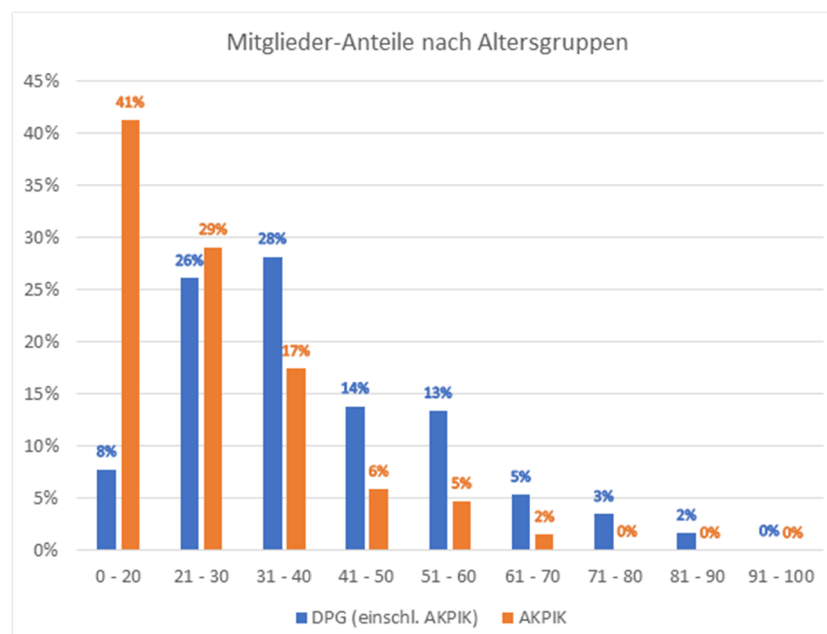
Die Aktivitäten des AKPIK richten sich genau auf diese Gruppen aus:

- In Gremien wie der Konferenz der Fachbereiche Physik wirken Mitglieder des AKPIK Vorstandes auf die Einbeziehung relevanter Themen (wie Maschinelles Lernen) ins Physikstudium hin. Zusammenarbeit besteht auch mit der Gesellschaft für Informatik, die unter dem Stichwort Data Literacy zur Förderung digitalen Kompetenzen an den Hochschulen in der Breite aufruft.

AKPIK auf Tagungen

von Martin Erdmann

- Die jährliche *Mitgliederversammlung* findet am Mittwoch 1.4.2020, 18:30 auf der DPG Frühjahrstagung Sektion Materie und Kosmos (SMuK) Universität Bonn Hauptgebäude statt.
<https://bonn20.dpg-tagungen.de>
- *The Science Cloud – Towards a Research Data Ecosystem for the next Generation of Data-intensive Experiments and Observatories*, WE-Heraeus-Seminar, 12 Jan - 15 Jan 2020, Physikzentrum Bad Honnef:



<https://www.we-heraeus-stiftung.de/veranstaltungen/seminare/2020/the-science-cloud-towards-a-research-data-ecosystem-for-the-next-generation-of-data-intensive-experiments-and-observatories>

- *Big Data Science in Astroparticle Research*, 17.-19.2.2020 im Super-C der RWTH-Aachen, Deep Learning, Hands-on Tutorial, Open Data – Software – Analysis, actual developments
<https://indico.scc.kit.edu/event/669>

Industriegespräche

von Andreas Fehlner

Anfang Dezember 2019 starten die DPG-Industriegespräche im Rhein/Neckar-Gebiet. Die Veranstaltungsreihe ist eine regionale Aktivität vom AIW und findet in Heidelberg unter Mitwirkung des AKPIK und der Heidelberg Laureate Forum Foundation statt.

Am Dienstag, 03. Dezember 2019, um 18:30 Uhr spricht Dr. Wolfgang Maier, Direktor Hardware Development des IBM R&D Labors in Böblingen über das Thema

„Künstliche Intelligenz und Quantum Computing - die Sicht der IBM Forschung & Entwicklung“

Ort der Veranstaltung ist das Mathematik-Informatik-Station (MAINS), Heidelberg Laureate Forum Foundation, Kurfürstenanlage 52, 69115 Heidelberg.

Weitere Details und die Möglichkeit sich zur Veranstaltung kostenlos anzumelden finden Sie unter <https://www.dpg-physik.de/vereinigungen/fachuebergreifend/ak/aIW/industriegespraecherheinneckar>

Für Rückfragen stehen ich Ihnen als Mitglied des Organisationsteams gerne unter der Adresse industriegespraecher-n@dpg-mail.de zur Verfügung.

Plattform Lernende Systeme (PLS)

Von Martin Erdmann und Katharina Morik

Die Plattform Lernende Systeme (PLS) ist ein Expertenkreis der forschenden Informatik mit Wirtschaftsunternehmen im Bereich der Künstlichen Intelligenz (KI) der vom BMBF gegründet und von der acatech verwaltet wird.

Weitere BMBF-Aktivitäten zur KI sind neben der langjährigen Förderung des Deutschen Forschungszentrums für Künstliche Intelligenz DFKI die Einrichtung von 4 Kompetenzzentren für Maschinelles Lernen und 2 Kompetenzzentren für Big Data. Ziel ist die Entwicklung einer KI zum Wohl des Menschen.

Im Unterschied zu anderen weltweiten Entwicklungen sollen hier Prüfsiegel entstehen, die z.B. für Verlässlichkeit der KI-Methoden stehen. Der Schutz persönlicher Daten orientiert sich an der europäischen Datenschutzverordnung, ethische Aspekte sollen während der Entwicklungen berücksichtigt werden (Ethics-by-Design), ein hohes Sicherheitsniveau soll unter anderem durch europäische Serversysteme und dezentrales Rechnen erreicht werden. Als wirtschaftliche Chance konzentriert man sich auf digitale Interaktionen zwischen Unternehmen (sogenanntes B2B, business-to-business), da der sogenannte B2C-Bereich (business-to-consumer) sehr weitgehend von Weltmarktführern erschlossen ist.

Sowohl die europäische Union als auch die Bundesregierung wollen die Entwicklung der KI mit Milliardenbeträgen fördern. Eine große Herausforderung ist dabei die Ausbildung von Nachwuchs im Bereich der KI sowohl für die Wissenschaft als auch für die Wirtschaft. Die Ausbildung an Universitäten muss durch neue Studiengänge erweitert, Professuren geschaffen werden.

Für diejenigen unter Ihnen, die noch am Beginn des Ausbildungs- und Berufsweges stehen, entstehen hier viele Chancen.
<https://www.plattform-lernende-systeme.de>

Erforschung von Universum und Materie (ErUM)

Von Martin Erdmann

Ein Teil der Physik-Forschung findet an Großforschungsanlagen statt: Teilchenbeschleuniger, Synchrotron-Strahlen, Neutronenquellen werden an Forschungszentren von Experten betrieben (CERN in Genf, DESY in Hamburg, ...).

Arbeitsgruppen aus den Universitäten reisen dorthin, um ihre Experimente zur Grundlagenforschung mit Bezug zu neuen Materialien und zum Verständnis unseres Universums durchzuführen.

Auf Initiative des Bundesministeriums für Bildung und Forschung (BMBF) haben Repräsentanten dieser Arbeitsgruppen mit insgesamt ca. 16.000 Mitarbeitern ein 40-seitiges Dokument zur Digitalisierung dieser Forschung für die kommenden 10 Jahre verfasst: Über die modernen Formen der Digitalisierung soll der Fortschritt in dieser Grundlagenforschung substantiell beschleunigt werden:

http://www.ketweb.de/stellungnahmen/e300611/Strategiepapier_ErUM-Data_Final_2019-04-29.pdf

Der Maßnahmenkatalog umfasst konkrete Empfehlungen zu allen 4 großen Themen, die oben im Begrüßungstext beschrieben sind: *Algorithmen, Computing, Datenmanagement, Ausbildung*. Unter dem Dach einer neuen Organisationsstruktur sollen die Maßnahmen gemeinsam umgesetzt werden.

Hier ergeben sich weitere tolle Chancen insbesondere für diejenigen unter Ihnen, die noch auf der Suche nach Ausgestaltung des beruflichen Wegs sind.

Nationale Forschungs-Daten-Infrastruktur NFDI

Von Martin Erdmann und Kai Polsterer

Die Wiederverwendung von Forschungsdaten zur Lösung neu auftauchender, wissenschaftlicher Fragestellungen soll auch Jahre nach der Beendigung eines Experiments noch möglich sein. Wünscht man sich. Aber das ist überhaupt nicht einfach. Zudem ist die Archivierung und Bereitstellung zwingend notwendig, um die Ergebnisse wissenschaftlicher Arbeiten reproduzieren zu können.

Wissenschaftler_innen in der Astronomie arbeiten bereits seit vielen Jahren an Infrastrukturen zum Umgang mit Daten verschiedenster Observatorien. Zum Beispiel können Sie in dem sogenannten internationalen virtuellen Observatorium nach Messungen von Objekten wie Galaxien und Sternen suchen. Die deutsche Beteiligung an diesem internationalen Projekt ist GAVO (German Astrophysical Virtual Observatory, <https://www.g-vo.org>).

In den Teilchenphysik-Experimenten am CERN steht dieser Prozess noch am Anfang. Hier haben die Detektoren 100 Millionen Sensoren in einer komplizierten Anordnung. Für die Interpretation der Detektormessungen benötigt man hochspezialisierte Software, die heute auf zertifizierter Computer-Hardware läuft. Jedes Experiment dokumentiert die aktuellen Ergebnisse in Journalpublikationen. Für die Zukunft: Speichert man Sensordaten oder besser bereits weiterverarbeitete Daten? Welche genau? Wie wurden die Daten weiterverarbeitet und analysiert? Wie kann man die Daten auch auf zukünftiger Computerhardware weiterhin verarbeiten?

In den hunderten von kleineren Laborexperimenten der Festkörperphysik sind die Datenmengen zwar viel geringer, aber die Komplexität liegt in den experimentellen

Bedingungen: Wie wurde die untersuchte Probe für das Experiment genau hergestellt, welche experimentellen Bedingungen herrschten während der Messung und was sind die Messunsicherheiten? In Journalpublikationen wird das alles dokumentiert, aber wie kann man all diese Informationen so zusammenstellen, dass eine effiziente Weiterverwendung möglich wird? Wie wird der Austausch zwischen unterschiedlichen Analysewerkzeugen ermöglicht?

Diese und ähnliche Fragestellungen sollen im Rahmen der NFDI-Initiative des Bundesministeriums für Bildung und Forschung (BMBF) von den Wissenschaftlern gelöst werden. Ca. 50 große Verbände haben sich in allen Wissenschaftsbereichen gebildet, um ihre Vorhaben zu beschreiben. Auch die DPG hat einen Verbund gebildet, damit alle Disziplinen der Physik an der NFDI-Initiative teilnehmen können. Die NFDI-Initiative steht noch ganz am Anfang, wie es weitergeht, werden wir Ihnen berichten.