



Universität
zu Köln



Kooperationsmöglichkeiten mit der Informatik der Uzk

Universität zu Köln | Prof. Dr.-Ing. Stefan Wesner |
Director ITCC, Chair „Parallel & Distributed Systems“

01

DIE JUNGE INFORMATIK AN DER UNIVERSITÄT ZU KÖLN

Dewey-Dezimalklassifikation (DDC) ist ein in Zehnergruppen gegliedertes System zur Ordnung von Wissen (die deutschen Übersetzungen stammen aus einem Projekt von 2002-2005)

Die tausend Klassen der dritten Ebene

510

510 Mathematik

511 Allgemeine mathematische Prinzipien

512 Algebra

000

000 Informatik

001 Wissenschaften

002 Das Buch

003 Systeme

004 Informatik

005 Computerprogrammierung, Computerprogramme & Daten

006 Spezielle Computerverfahren

519 Wahrscheinlichkeiten, angewandte Mathematik

Die Informatik ist ein hochdynamisches Umfeld.

Eine junge Informatik ist damit
insbesondere eine moderne Informatik.

Die Ausrichtung der Kölner Informatik ist bewusst auf ausgewählte Kernbereiche fokussiert

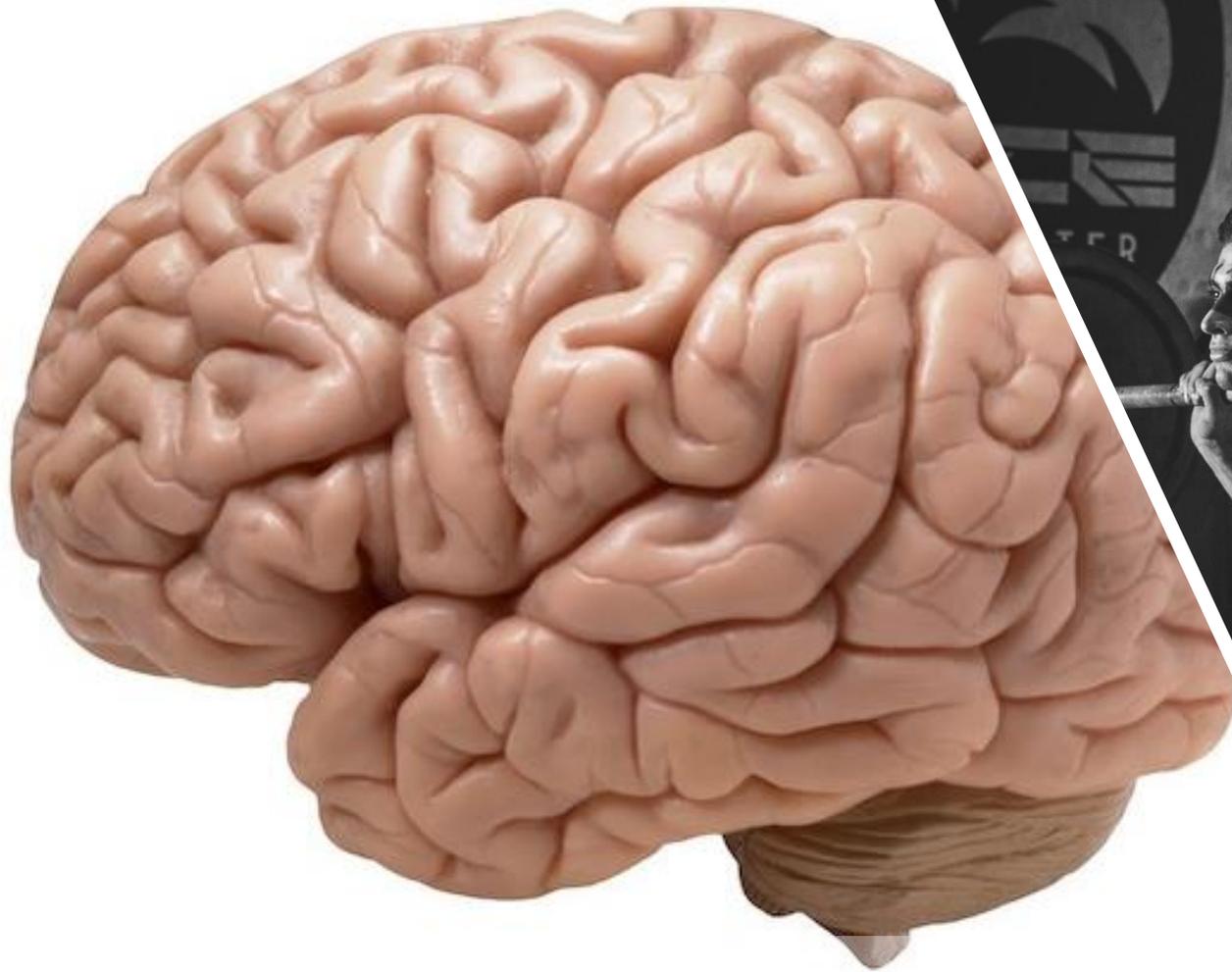
- Der gewählte thematische Schwerpunkt der Informatik an der UzK orientiert sich neben den beständigeren theoretischen Grundlagen an Themen mit besonderer aktueller Relevanz.
- **Algorithmen und Theorie**
 - Effizienz von Algorithmen
 - Quanteninformation
- **Engineering und Software-intensive Systeme**
 - Immer kürzer werdende Innovationszyklen für gleichzeitig immer komplexere Systeme
 - Softwarequalität, -robustheit und -resilienz
 - Sicherheit von Softwaresystemen
- **Artificial Intelligence & Visual Analytics**
 - Neue Methoden, Effiziente Anwendung von bestehenden Methoden des maschinellen Lernens
 - Umgang mit sehr großen, sehr komplexen und sehr schnell veränderlichen Daten
 - Effiziente Darstellung und Bewertung von unüberschaubaren Datenmengen
- **Wiss. Rechnen und HPC**
 - Nutzung sehr großer Rechensysteme, um bestehende Probleme schneller und größere Probleme berechenbar zu machen
 - Quantencomputing

02

INDUSTRIE UND UNI?

**DAS SIND DOCH ALLES
THEORETIKER!**





Brain meets Brawn

Notwendige Elemente einer Kooperation



+



+



+



Expertise

Wissen zum
Stand der
Technik und
mehr

Direkt verfügbares
Personal

Studierende,
Doktoranden,
Post-Docs

Kaffee

Wenn das Problem
spannend ist, wird auch
länger daran gearbeitet,
und oft beginnt das
Brainstorming mit dem
Industriepartner mit
einem Kaffee

Infrastruktur

**Viele Fragestellungen
gerade aus dem
Bereich ML erfordern
Zugang zur
Infrastruktur**

Alles Icons von [Icons8](#)

Unterschiedliche Kooperationsmodelle sind möglich

- Untersuchung einer Fragestellung im Rahmen einer Abschlussarbeit oder als Teil eines Softwareprojektes durch **Studierende**
- Gemeinsames Interesse an einer Fragestellung, Austausch von Expertise und zum Beispiel Bereitstellung von Daten/Informationen. Gerade **Doktorand:innen** suchen nach konkreten Fragestellungen zur Validierung ihrer Lösungsansätze und Ideen.
- Zusammenarbeit im Kontext spezifischer Ausschreibungen wie „Zentrales Innovationsprogramm Mittelstand“, KMU-innovativ oder anderer Ausschreibungen auf nationaler und europäischer Ebene. **Mehrwert ist die Expertise des Lehrstuhls und dessen internationales Netzwerk (kommt aber mit etwas Bürokratie – je nach Mittelgeber)**
- Direkte Kooperation/Auftragsforschung (**fokussiert, Nutzung bestehender Vorarbeiten und Ressourcen – aber kostenintensiv**)
- Im Rahmen einer typischen Kooperation kommen mehrere der obigen Modelle oder Mischformen zum Einsatz

03

BEISPIELE

Beispiel: Synthetisierung von Zeitreihendaten



>100 TB Zeitreihendaten (IP Flows eines Netzbetreibers) mit sensiblen Daten sollen für die Validierung von Produkten verwendet werden, können aber nicht weitergegeben werden



Lösungsansatz: Nutzung von Generative Adversarial Networks (GANs) zur Synthetisierung von Time Series Daten mit gleichen Eigenschaften

Kooperationsmodell: Beispieldaten und Use Case für ein Promotionsvorhaben
Validierung der Methode im Kontext von realer Fragestellung
Prototypische Implementierung der Methode unter Open Source für Firma verfügbar



Nutzung von lokaler GPGPU Infrastruktur für die Entwicklung und Validierung des Lösungsansatzes

Beispiel 2a: Skalierbare Datenbanken in der Cloud



Der Übergang von einer zentralen Datenbankinstanz zu einer verteilten Datenbank in der Cloud liefert nicht den erwarteten Leistungsgewinn



Lösungsansatz: Systematische Untersuchung des Skalierungsverhalten und der Leistung bei verschiedenen Cloud Anbietern

Kooperationsmodell: Direkte Finanzierung eines Doktoranden für 3 Monate, Gemeinsame Erarbeitung einer Testumgebung und dem Deployment auf der lokalen Cloud Infrastruktur des Lehrstuhls, Systematisches Messen der Antwortzeiten unter synthetischer Last und künstlich herbeigeführter Fehlersituationen. Gesamtkosten ca. 30T€

Vorschlag eines Workarounds zur kurzfristigen Lösung



Nutzung einer lokalen Cloud Infrastruktur auf Basis von OpenStack

Vorarbeit: Software Framework zum systematischen Messen von Leistung in der Cloud aus einem EU-Forschungsprojekt

Beispiel 2b: Skalierbare Datenbanken in der Cloud



Neben einer Ausweitung der Untersuchung aus dem Vorprojekt auf mehr Datenbankprodukte, mehr Fehlerfälle und weitere Cloud-Anbieter sollte untersucht werden, ob auch Vorhersagen für Leistungsdaten ohne Deployment realisiert werden können.

(Testläufe und Benchmarks in der Cloud sind teuer und zeitaufwändig.)



Lösungsansatz: Erarbeitung einer neuartigen Methodik zur Erkennung und Messung von Problemen beim Betrieb verteilter Infrastruktur in Cloud-Umgebungen und, darauf aufbauend, zur Vorhersage von Leistungs- und Verlässlichkeit verschiedener alternativer Lösungsansätze für die verteilte Datenhaltung (=Promotionsthema)

Kooperationsmodell: Direkte Finanzierung eines Doktoranden für 24 Monate

(Ergebnis nach weiteren 2 Jahren: das Start-Up benchAnt, <https://benchant.com>)



Nutzung einer lokalen Cloud-Infrastruktur auf Basis von OpenStack

Umfangreiche Validierung auch bei öffentlichen Cloud-Anbietern

Alles Icons von [Icons8](https://icons8.com)

Beispiel 3: KI-gestützte Softwareentwicklung



Integration von KI-gestützten Methoden bei der Entwicklung in ein bestehendes kommerzielles Werkzeug zur Modellierung von Geschäftsprozessen und Systemintegration



Lösungsansatz: Realisierung eines gemeinsamen Projektantrags mit weiteren Partnern aus Wissenschaft und Industrie

Antrag im Horizon Europe Förderprogramm derzeit in Vorbereitung

Kooperationsmodell: Externe Finanzierung notwendig. Kooperation mit Partnern aus den jeweiligen Netzwerken in einem gemeinsamen Projekt mit einer Laufzeit von 36 Monaten.

Erfolgschancen derzeit 1:10 – 1:30 je nach Ausschreibung



Verfügbare Infrastruktur wird bei Anträgen für Partner als „in-kind“ eingebracht

Alles Icons von [Icons8](#)

04

TAKE-HOME MESSAGES

Algorithmen und Theorie



Algorithm Theory
(2024)



Efficient Algorithms
(2024)



Alg. Datenanalyse
(Sohler)

Engineering Software-Intensiver Systeme



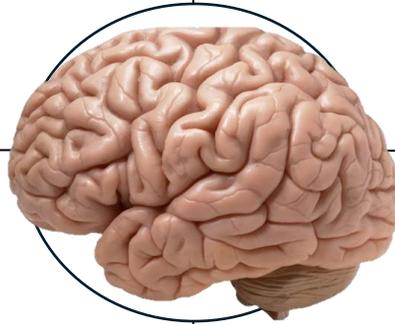
Software Engineering
(Vogelsang)



IT Security
(2024)



Softwaretechnologie
(Felderer)



Artificial Intelligence & Visual Analytics



Artificial Intelligence
(Frahling)



Visual Analytics
(Landesberger)

Wissenschaftliches Rechnen & HPC



Quantencomputing
(2024)



Data &
Information Science
(Bojchevski)



Parallel &
Distributed Systems
(Wesner)

+PhD Candidates and
Post-Docs in all areas!



Machine Learning
2024



SciVis
(Zellmann)



Computational Earth
System Science
(Schultz)

Prof. Dr.-Ing. Stefan Wesner, Director ITCC &
Chair "Parallel & Distributed Systems"

Ressourcen

- **Infrastruktur**

- zur Speicherung sehr großer Datenmenge (Petabytes = 10^{15} byte = 1 Million Gigabyte)
- Durchführung massiver Datenanalysen auf GPGPUs (ab März 2024)
- Dezierte Infrastruktur zur Validierung und Leistungsanalyse von Software
- Visualisierungsinfrastruktur (unter anderem 5 Seiten CAVE Umgebung)
- Zugriff auf neuartige und experimentelle Infrastruktur (Hardware, Software) durch Kooperation mit verschiedenen Anbietern

- **Netzwerk**

- Kooperation mit Wissenschaftlern aus den Anwendungsbereichen der UzK, als auch mit internationalen Partnern



Take-Home Messages

- **Die Informatik an der Universität zu Köln**
 - Ist fachlich klar fokussiert auf vier sehr aktuelle Themenbereiche
 - Hat eine sehr enge Anbindung an die Anwendung in verschiedenen Disziplinen von Medizin über Naturwissenschaften bis hin zu den Geisteswissenschaften
- Die Kooperation mit einer Universität ist nicht auf die Professorin/den Professor beschränkt, denn es gibt immer ein ganzes Team und eine umfangreiche Infrastruktur.
- Es gibt im Team der Informatik@UzK umfangreiche Erfahrungen bei der Zusammenarbeit mit der Industrie
- Es gibt viele Möglichkeiten der Kooperation, von gemeinsamen Abschlussarbeiten oder Softwareprojekten bis hin zu gemeinsamen mehrjährigen Forschungsprojekten



UNIVERSITY
OF COLOGNE

VIELEN DANK FÜR DIE AUFMERKSAMKEIT

FRAGEN?

wesner@uni-koeln.de

info@cs.uni-koeln.de