

Prof. Dr. habil. Nihat Ay

Data Science Foundations

FORSCHUNGSKONZEPT

Das Ziel

Das Institut für Data Science Foundations verfolgt das Ziel, **allgemeine Prinzipien und Methoden der datengetriebenen Wissensextraktion** zu entwickeln. Zentral hierbei ist die konzeptionelle und mathematische Herangehensweise.

Das Konzept

Daten resultieren aus einem interaktiven Prozess, der durch die **sensomotorische Schleife** beschrieben wird (Abb. 1). Wissen als Grundlage für informierte Entscheidungen ist grundsätzlich im Kontext dieser Schleife zu verstehen.

Die Methoden

Wir verwenden ein **breites Spektrum an abstrakten Strukturen und mathematischen Methoden**, die auch in der Lehre vermittelt werden (siehe Schwerpunkt Lehre). Einen Schwerpunkt bilden hierbei geometrische Methoden zur Beschreibung und Optimierung von Lernprozessen.

Eine Anwendung: Optimale Steuerung cyber-physischer Systeme

Wir werden zur **optimalen Steuerung cyber-physischer Systeme** beitragen. Diese stellen ein zentrales und profilbildendes Thema der TUHH dar (Abb. 2).

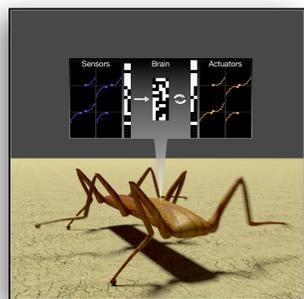


Abb. 2: Steuerung eines cyber-physischen Systems mit einem neuronalen Netz. Ein Beispiel für „cheap control“.

Abb. 3: Das Motto der TUHH im Kreis.



AUSGEWÄHLTE PROJEKTE

Informationsgeometrische Strukturen in Data Science

Die Informationsgeometrie hat zum Ziel, **natürliche geometrische Strukturen** zu identifizieren. Diese Strukturen erlauben es, Lernprozesse zu optimieren und damit Wissen effizienter aus den Daten zu extrahieren (Abb. 4).

Verkörperte Intelligenz

Die Wissensextraktion findet im Kontext der sensomotorischen Schleife statt. Das Gebiet der **verkörperten Intelligenz** stellt hierbei wesentliche Konzepte bereit, z. B. „cheap control“ und „morphological computation“ (Abb. 2 & 5).

Komplexität und Kausalität in Daten

Struktur und Komplexität der Daten entsteht durch kausale Wechselwirkung. Das Verständnis von **Kausalität ist notwendig, um „actionable knowledge“ zu extrahieren** (Abb. 6). Korrelationale Struktur alleine ist unzureichend.

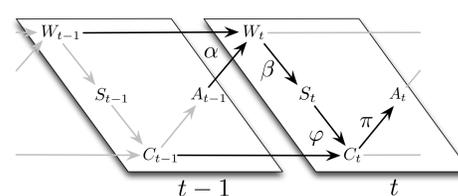
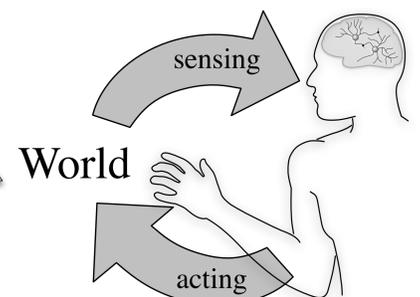


Abb. 1: Das kausale Diagramm der sensomotorischen Schleife.



SCHWERPUNKT LEHRE

Vorlesungen & Seminare zu Themen im Bereich Data Science

Neuronale Netze, maschinelles Lernen, graphische Modelle, Informationstheorie, geometrische Methoden in der Statistik; **Bachelor-Studiengang Data Science** (Master-Studiengang in der Planung).

Das Ziel

1. Studierende der TUHH dazu befähigen, die **bestehenden Data-Science-Methoden anzuwenden, diese weiter zu optimieren und neuartige Methoden zu entwickeln**.
2. Motto „**Technik für die Menschen**“ der TUHH allgemeiner auffassen und insbesondere im Kreis als „**Menschen für die Technik**“ lesen (Abb. 3).

Das Konzept

1. Zentrale **Grundideen vermitteln** und damit **einen Kontext schaffen für komplexere Zusammenhänge**.
2. Wichtige **mathematische Strukturen vermitteln**; dabei keine Reduktion der Inhalte auf „Kochrezepte“.
3. Mit Hilfe von Übungen, Praktika und algorithmischen Implementierungen die **Inhalte konsolidieren**.

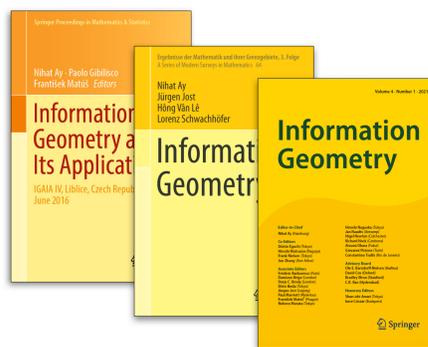


Abb. 4: Forschung & internationale Vernetzung durch Konferenzen und Herausgeberschaften im Bereich der Informationsgeometrie.



The Active Self

SPP funded by DFG

Abb. 5: DFG-geförderte Projekte im Rahmen der Schwerpunktprogramme „Autonomous Learning“ (DFG, SPP 1527) und „The Active Self“ (DFG, SPP 2134), 2012-2014, 2015-2018, 2018-2021, 2021-2024.



Mathematics for Multilevel Anticipatory Complex Systems

Abb. 6: EU-gefördertes FET-Projekt (FP7-ICT) mit 6 Partnerinstitutionen aus 4 Ländern, 2012-2015.

