

ÖFFENTLICHER VORTRAG

DR. CORDELIA SCHMID

Forschungsdirektorin am Institut National de Recherche
en Informatique et en Automatique (INRIA)



SEMINAR, exklusiv für Studierende
und KIT-Angehörige am 23.10.2024

AM 22.10.2024

**HEINRICH-HERTZ
GASTPROFESSUR 2024**

Das Karlsruher Institut für Technologie und der
KIT Freundeskreis und Fördergesellschaft e.V. verleihen
am Dienstag, 22. Oktober 2024 die

HEINRICH-HERTZ-GASTPROFESSUR 2024 **an Dr. Cordelia Schmid**

VORTRAG AM 22. OKTOBER 2024

Campus Süd, Johann-Gottfried-Tulla-Hörsaal,
Geb. 11.40, Englerstraße 11, 76131 Karlsruhe

■ **17:00 – 18:30 Uhr**

Begrüßung und Preisverleihung durch Professor Dr.
Jan S. Hesthaven, Präsident des KIT, und Dr. Jürgen
Spitzer, Vorsitzender des Vorstands des KIT Freundes-
kreis und Fördergesellschaft e.V.

Vortrag von Dr. Cordelia Schmid:

„Multimodales Verstehen von Video-Datenströmen
und Bild-Sprache gesteuerte Robotik“

STUDIARENDESEMINAR AM 23. OKTOBER 2024

Campus Süd, Gastdozentenhaus, Geb. 01.52,
Dürer-Saal, Engesserstraße 3, 76131 Karlsruhe

■ **10:00 – 11:30 Uhr**

„Multi-modal reasoning for video understanding
and 3D scene generation“

UM ANMELDUNG WIRD GEBETEN.

Bitte nutzen Sie dafür den folgenden Registrierungslink:
<https://www.cse.kit.edu/4042.php>.
Anmeldeschluss ist der 15. Oktober 2024.

VORTRAG

Im Vortrag werden Fortschritte beim Lernen aus großen multimodalen Video-Datenströmen vorgestellt. Konkret wird ein effizientes Modell zur dichten Video-Datenerfassung erläutert, das als Eingabe Videodaten und Sprache verwendet und sowohl zeitliche Ereignisse als auch Textbeschreibungen gleichzeitig vorhersagen kann. Anschließend wird eine Methode zur Beantwortung von Fragen zu Video-Daten und Bildsequenzen vorgestellt, die auf einem durch Suchaufrufe erweiterten visuellen Sprachmodell basiert. Das System lernt, umfangreiches Weltwissen zu kodieren und abzurufen, um wissensintensive Abfragen zu beantworten. Dieser Ansatz erzielt effektive Ergebnisse bei der Beantwortung visueller Eindrücke und Fragen sowie der Bilduntertitelung.

Im zweiten Teil des Vortrags werden Arbeiten zur Unterstützung visuell geführter Navigation und Roboteranipulation anhand von Sprachanweisungen vorgestellt. Erweiterte Vision-Language-Transformer, die Handlungshistorien integrieren und Aktionen vorhersagen, kommen zum Einsatz. Der History Aware Multimodal Transformer (HAMT) übertrifft den Stand der Technik bei verschiedenen Benchmarks für Bild-Sprache-Navigation. Weitere Verbesserungen lassen sich durch Integration von Karteninformationen erreichen. Am Beispiel eines Roboters wird die Objekt-Zielnavigation in realen Umweltszenarien demonstriert. Abschließend wird gezeigt, dass Transformerbasierte Ansätze auch zur Unterstützung Roboterbasierter Handhabungen effektiv sind, insbesondere durch visuelle 3D-Darstellungen. Ein UR5-Roboter-Arm zeigt die hohe Performanz bei konkreten Anwendungen.

BIOGRAFIE

Dr. Cordelia Schmid ist eine deutsche Informatikerin. Sie erzielte bedeutende Fortschritte im Bereich der künstlichen Intelligenz, speziell im Bereich „sehender“ Maschinen und der Erkennung von Mustern, Gesichtern und Handlungen in Fotos und Videos. Sie studierte Informatik an der damaligen Universität Karlsruhe und promovierte in Informatik am Institut National Polytechnique de Grenoble (INPG). Ihre Dissertation über „Lokale Grauwertinvarianten für Bildabgleich und -abruf“ wurde 1996 als beste Dissertation ausgezeichnet. 2001 habilitierte sie sich mit ihrer Arbeit „Vom Bildabgleich zum Lernen visueller Modelle“. Cordelia Schmid war von 1996 bis 1997 Postdoktorandin in der Robotics Research Group der Universität Oxford. Seit 1997 hat sie eine feste Forschungsstelle bei INRIA inne, wo sie Forschungsleiterin ist. Cordelia Schmid ist Mitglied der Deutschen Nationalen Akademie der Wissenschaften Leopoldina, sie ist IEEE Fellow und ELLIS Fellow. Für ihre grundlegenden Beiträge zum effektiven Maschinensehen wurde sie 2006, 2014 und 2016 mit dem Longuet-Higgins-Preis und 2018 mit dem Koenderink-Preis ausgezeichnet. Sie erhielt 2013 einen ERC Advanced Grant, 2015 den Humboldt-Forschungspreis, 2020 den Royal Society Milner Award, 2021 den PAMI Distinguished Researcher Award und 2023 den Körber European Science Award sowie den Europäischen Erfinderpreis 2024 in der Kategorie „Forschung“.



Foto: Marcus Gloger | Körber-Stiftung

LAGEPLAN | ANFAHRT

1 | VORTRAG

Campus Süd, Johann-Gottfried-Tulla-Hörsaal,
Geb. 11.40, Englerstraße 11, 76131 Karlsruhe

2 | SEMINAR

Campus Süd, Gastdozentenhaus Geb. 01.52,
Dürer-Saal, Engesserstraße 3, 76131 Karlsruhe

Auf dem Gelände des **KIT Campus Süd** stehen **nur begrenzt Parkplätze** zur Verfügung. Bitte nutzen Sie die Möglichkeit der Anreise mit öffentlichen Verkehrsmitteln, Zielhaltestelle Kronenplatz.

ANMELDUNG

<https://www.cse.kit.edu/4042.php>.

Anmeldeschluss ist der 15. Oktober 2024.



GREEN
EVENT
BW 

Kontakt

Telefon: +49 721 608-22011

E-Mail: event@kit.edu

Herausgegeben von

Karlsruher Institut für Technologie (KIT)
Präsident Prof. Dr. Jan S. Hesthaven
Kaiserstraße 12
76131 Karlsruhe
www.kit.edu