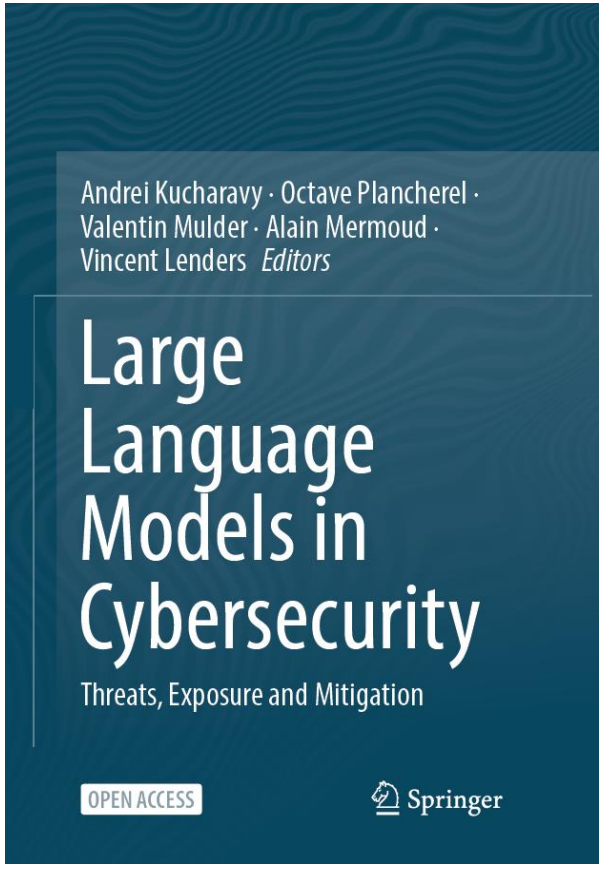
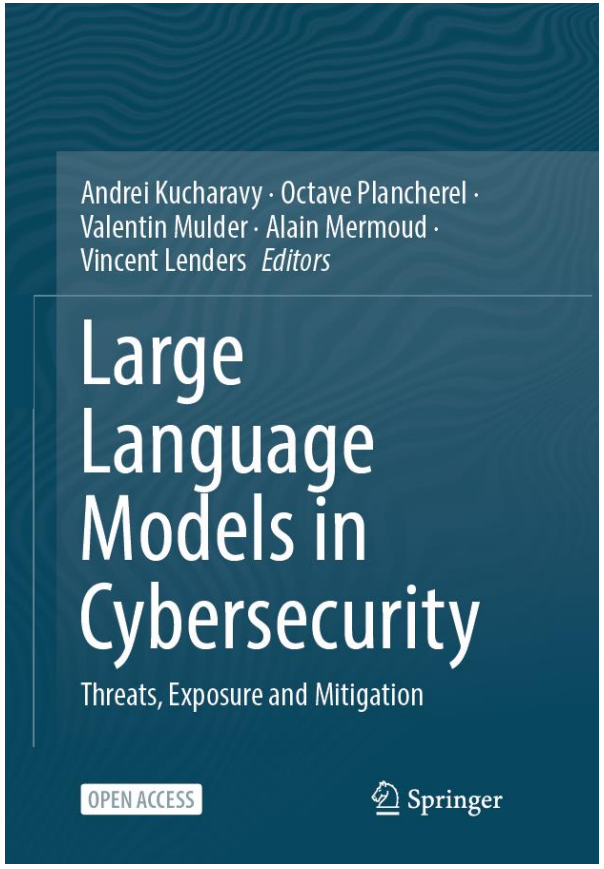
Wissenschaftliches Interesse der Studie



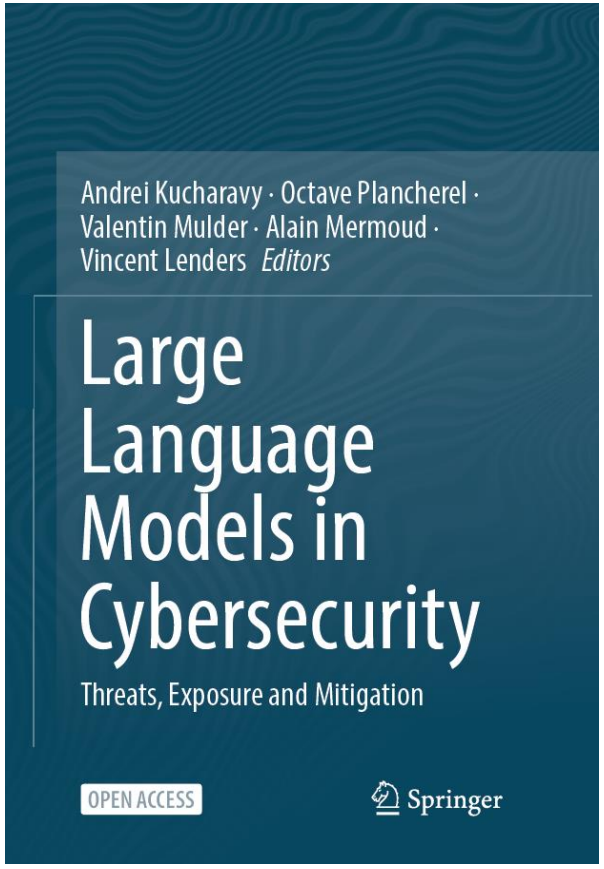
- Im Einklang mit der Forschung des Gen Learning Center
  - Angewandte Forschung zur Sicherheit und Fairness von generativem ML
- Wissenschaftlicher Stand der Technik
  - Kritische Informationen für Entscheidungsträger
  - Umfassende Einführung für technische Experten
  - Neuartige Forschung

Weltklasseniveau -Zusammenarbeit



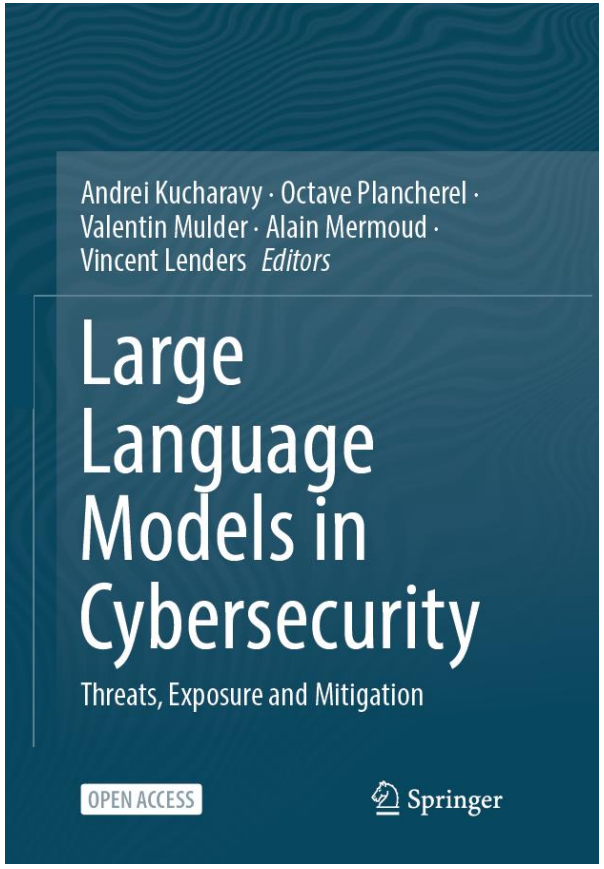
- Akademische Forschung :
  - ZHAW, EPFL, UNIL, UNIFR
  - Université de Paris, University of Victoria BC
  - Swiss AI Initative
- Angewandte Forschung in Industrien:
  - Effixis, Kudelski
  - IBM/Sophos/Censys

Wissenschaftlicher Stand der Technik



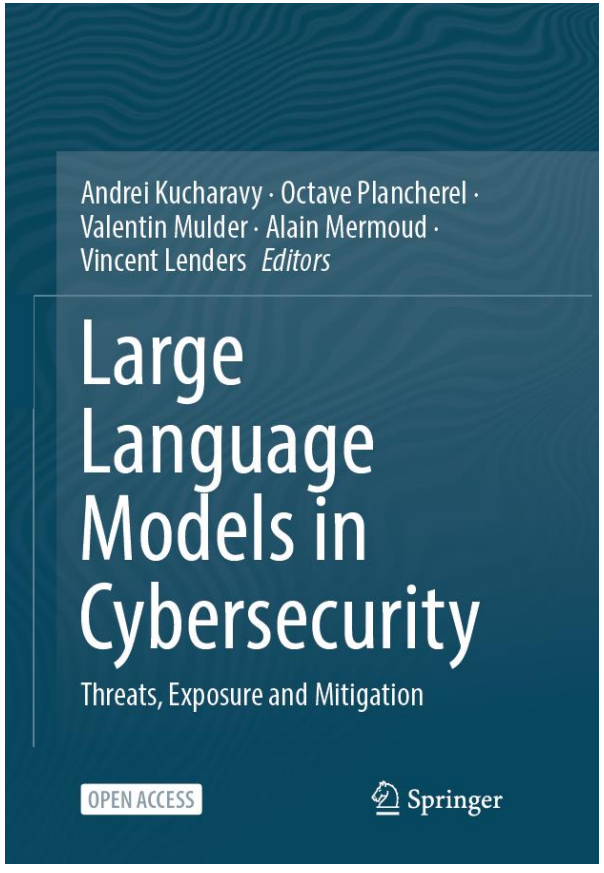
- Phishing Demonstration
- Deep Web-Indizierung Demo
- LM-Risikoversicherung Prognose
- Schutz der Privatsphäre beim Modelltraining
- gegensätzliche Beispiele für LLMs Abmilderung
- universelle Grundsätze der Normen

Immer noch aktuell



- Grundlagenforschung
  - Dauerhafte Konzepte
- Wir schätzen, dass sie bis ~2028 SotA bleiben wird.
  - Wissenschaftliche Begleitung bevor
  - Konzipiert als lehrbuchmäßige Einführung in das Fachgebiet nach diesem

Schweizer Studie Weltweiter Wirkung



The EU has started work on its Cyber Resilience Act, with provisional agreement having been reached in November 2023 between the Council and the European Parliament on security requirements for digital products. For these negotiations to lead to effective and future-proof legislation, thorough, scientifically based, analysis is an absolute requirement.

This book provides, in a most timely manner, such analysis. It provides cyber-security practitioners with tools to mitigate LLMs threats, LLM developers with an awareness of their models' vulnerabilities, misuse and potential mitigations and policymakers with a clear and convincing explanation of risks of LLMs in cybersecurity. It also gives a convincing indication of where science is going in the field of cybersecurity. This important academic work is therefore a must-read not only for researchers, cybersecurity specialists, law enforcement and defence officials but for anyone who takes an interest in the rapid and transformative changes generated by breakthrough technologies like LLMs.

Consultant, Member Advisory Board ALLAI
Former Director Council of Europe
Strasbourg, France
December 2023

Jan Kleijssen