

Open Source, KI und digitale Regulierung: Zwischen Offenheit, Innovation und rechtlicher Verantwortung

Melissa Irtel

JUN Legal GmbH

Herbstakademie 2025

1 Einleitung



- ▶ Zunehmende Bedeutung von Open Source KI-Initiativen, z.B. DeepSeek
 - ▶ Transparenz, Offenheit und weltweiter kollaborativer Austausch als Triebfeder technischer Innovation



- ▶ Rechtliche Herausforderungen, insbesondere beim Training mit frei verfügbarem Source Code
- ▶ Hinreichende Regulierung?

2 Grundlagen und Definitionsansätze

Open Source

= Free and Open Source Software

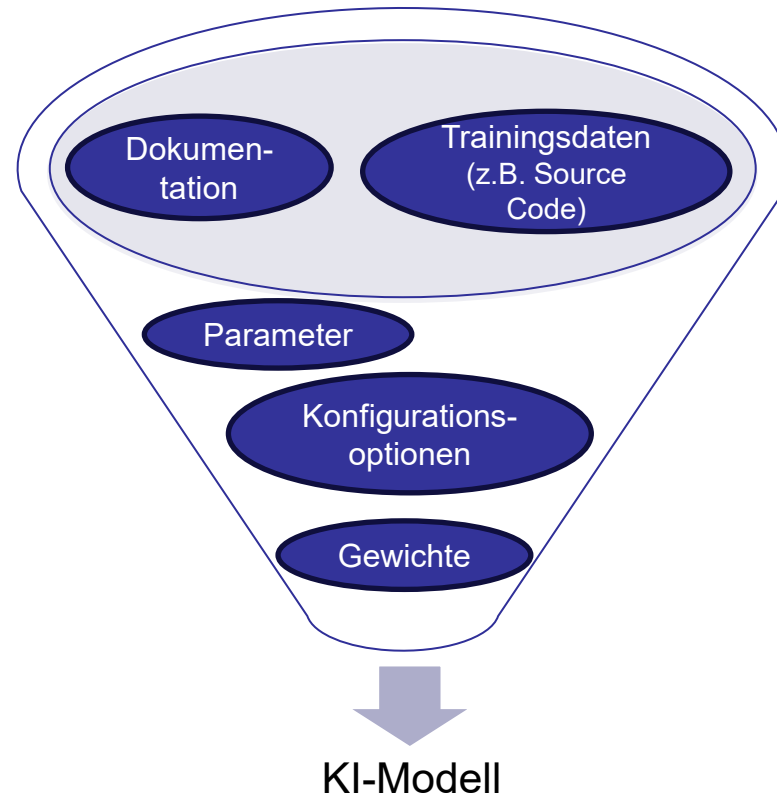
- Von jedermann frei und kostenlos verwendbar unter Einhaltung der Lizenzbedingungen
- Source Code zugänglich
- Lizenzen technologieneutral und nicht produktspezifisch
- Keine Diskriminierung/Nutzungsbeschränkung

Open Source KI

noch keine einheitliche Definition

Open Source Definition reicht nicht aus

2 Grundlagen und Definitionsansätze



Vereinfachte Darstellung von Bestandteilen eines KI-Modells
→ verschiedene Grade an Offenheit möglich

2 Grundlagen und Definitionsansätze

Open Source AI Definition der Open Source Initiative (OSI)
= OSAID-1.0

**Ein kostenloses KI-System gilt als Open Source KI, wenn
Nutzer die Freiheit haben**

- das KI-System für jeden Zweck zu nutzen, ohne um Erlaubnis fragen zu müssen
- zu untersuchen, wie das KI-System funktioniert und zu verstehen, wie es zu den Ergebnissen kommt
- das KI-System für jeden Zweck zu verwenden, u.a. auch den Output zu verändern, und
- das KI-System mit anderen zu teilen

und wenn Code, Daten und Parameter kostenlos zur Verfügung gestellt werden.

2 Grundlagen und Definitionsansätze

- ▶ Bislang noch kein Standard bzgl. Lizenzierung von Open Source KI
- ▶ Erste Ansätze zur „offenen“ Lizenzierung:

RAIL	OpenMDW-1.0
<i>BigScience</i>	<i>Linux Foundation</i>
u.a. Lizenz für BLOOM	im Entwurfsstadium, Lizenz für ByteDance's Seed-X-7B
permissive Lizenz mit Nutzungsbeschränkungen	vollständig permissiv, ohne Nutzungsbeschränkungen

3 Aktuelle Regulierung von Open Source KI

KI-Verordnung grds. nicht anwendbar für KI-Systeme, die **unter freien und quell-offenen Lizenzen** bereitgestellt werden (Art. 2 Abs. 12 KI-VO), es sei denn, es handelt sich um:

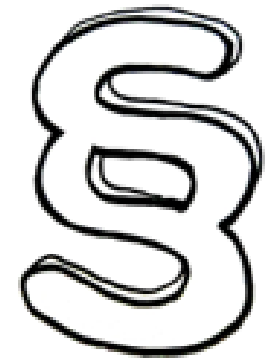
Rückausnahme 1:
Hochrisiko-KI-
Systeme,
Art. 2 Abs. 12
Hs. 2 KI-VO,

Rückausnahme 2:
verbotene KI-
Systeme,
Art. 5 KI-VO,

Rückausnahme 3:
KI-Systeme nach
Art. 50 KI-VO,

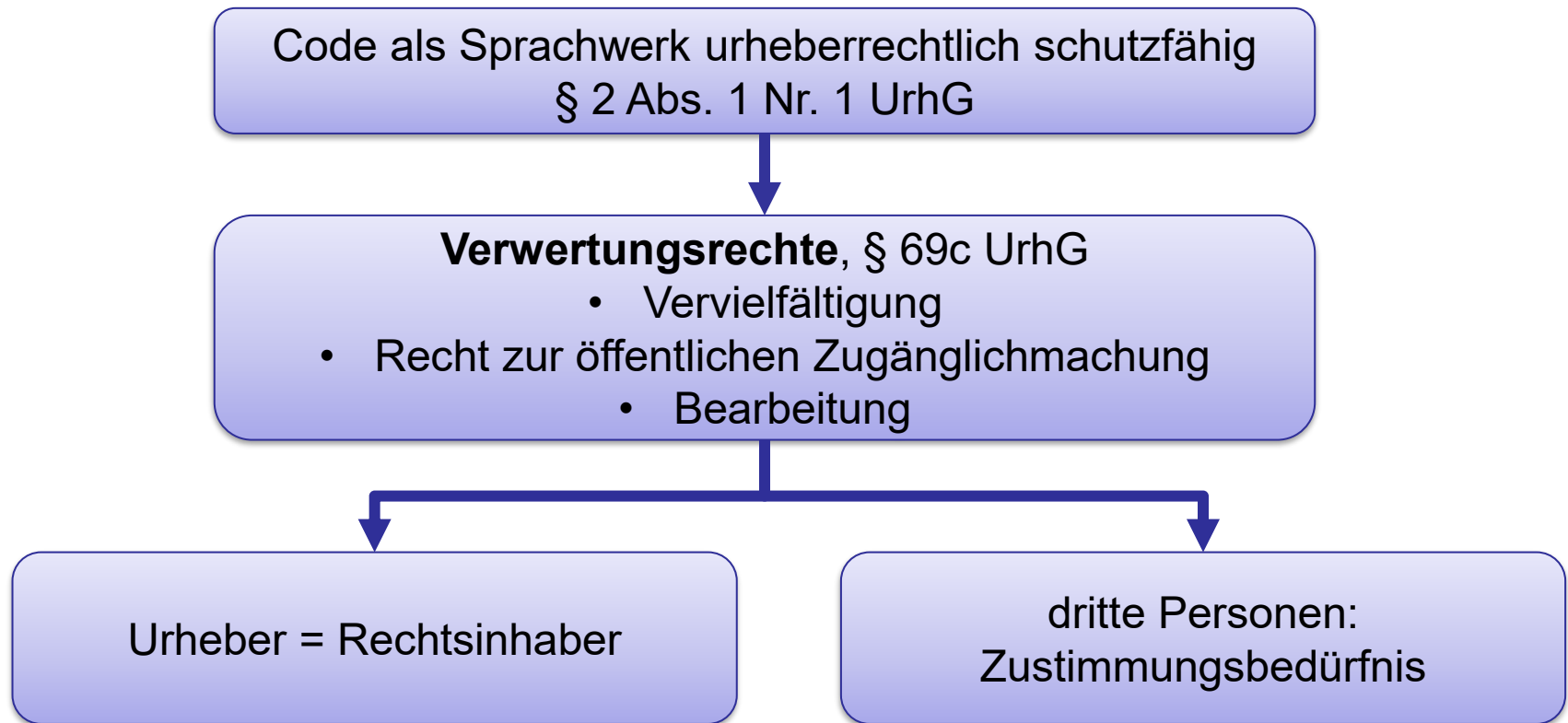
Rückausnahme 4:
kommerzielle KI-
Systeme
(*Umkehrschluss*),

- ▶ Keine Definition von Open Source KI
- ▶ Bezugnahme auf „Quelloffenheit“ unzureichend
- ▶ Anreiz für „Open Washing“
 - einheitliche Definition erforderlich



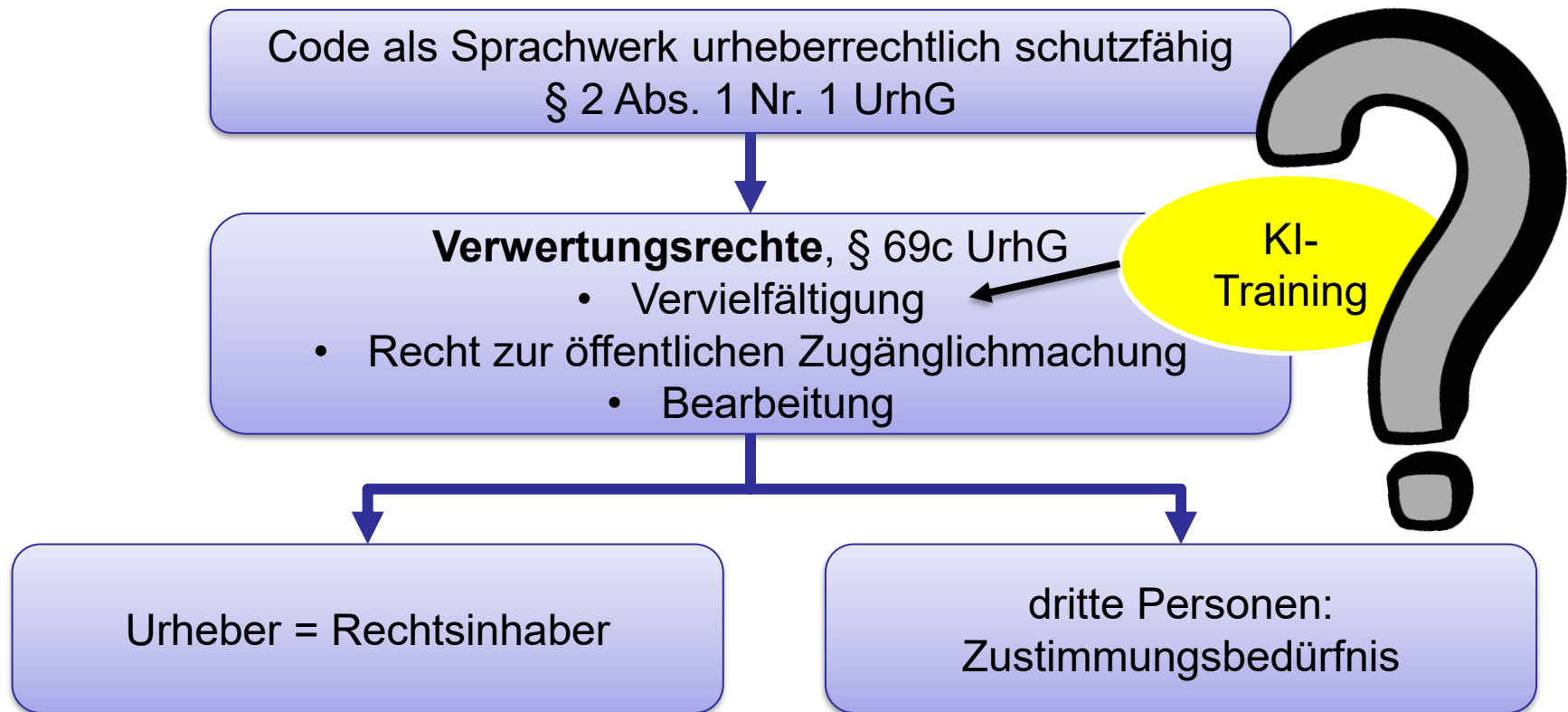
4 Rechtliche Herausforderungen im Zusammenspiel von Open Source und KI

Urheberrecht und KI-Training mit Open Source Software



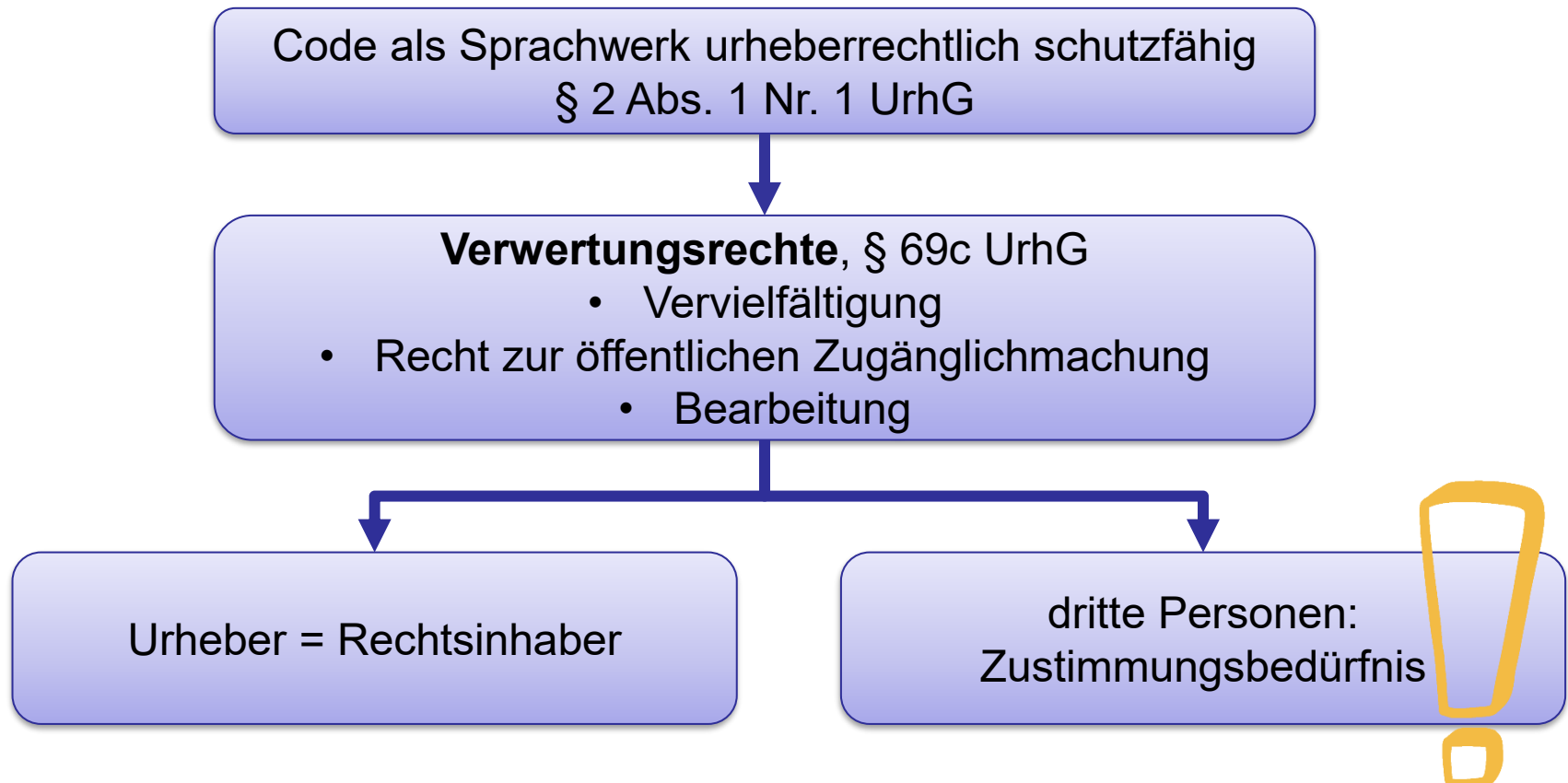
4 Rechtliche Herausforderungen im Zusammenspiel von Open Source und KI

Urheberrecht und KI-Training mit Open Source Software



4 Rechtliche Herausforderungen im Zusammenspiel von Open Source und KI

Urheberrecht und KI-Training mit Open Source Software



4 Rechtliche Herausforderungen im Zusammenspiel von Open Source und KI

Urheberrecht und KI-Training mit Open Source Software



fehlende Zustimmung des Rechtsinhabers

bei gesetzlicher Schranke keine Zustimmung erforderlich

Text- und Data-Mining, § 60d UrhG / § 44b Abs. 2 UrhG
→ anwendbar auf KI-Training?



(+) vgl. LG Hamburg, Urt. v. 27.09.2024 – 310 O 227/13
(LAION)

→ grds. Zulässigkeit des KI-Trainings mit FOSS

4 Rechtliche Herausforderungen im Zusammenspiel von Open Source und KI

Urheberrecht und KI-Training mit Open Source Software



ausnahmsweise Unzulässigkeit bei strengen Copyleft-Lizenzen, z.B. GPL-2.0

§4. Sie dürfen das Programm nicht vervielfältigen, verändern, weiter lizenzieren oder verbreiten, sofern es nicht durch diese Lizenz ausdrücklich gestattet ist. Jeder anderweitige Versuch der Vervielfältigung, Modifizierung, Weiterlizenzierung und Verbreitung ist nichtig und beendet automatisch Ihre Rechte unter dieser Lizenz. Jedoch werden die Lizenzen Dritter, die von Ihnen Kopien oder Rechte unter dieser Lizenz erhalten haben, nicht beendet, solange diese die Lizenz voll anerkennen und befolgen.

4 Rechtliche Herausforderungen im Zusammenspiel von Open Source und KI

Output nach KI-Training mit FOSS

Urheberrechtliche Schutzfähigkeit bei

- ▶ hinreichender Schöpfungshöhe (§ 2 Abs. 2 UrhG)
- ▶ menschlicher, geistiger Schöpfung

✗ Entwickler des KI-Systems
? Benutzer (Prompter)
? Urheber der Trainingsdaten



- ▶ andernfalls: gemeinfrei

4 Rechtliche Herausforderungen im Zusammenspiel von Open Source und KI

Output nach KI-Training mit FOSS

Praktische Risiken der Gemeinfreiheit des Outputs

- ▶ Potenzielle Sicherheitslücken und Fehler
- ▶ Mangels Urheberschaft keine Veröffentlichung als Open Source Code
 - daher auch keine Vorteile von FOSS
 - Verbesserung durch kollaboratives Zusammenwirken
 - Rechtssicherheit
 - Anreiz zur Innovation

→ **Anpassung des Urheberrechts** z.B. durch neues Leistungsschutzrecht für KI-generierten Code?

4 Rechtliche Herausforderungen im Zusammenspiel von Open Source und KI

Output nach KI-Training mit FOSS

Urheberrechtsverletzung durch **Vervielfältigung** im Output?

- ▶ i.d.R. keine Reproduktion des Trainingsmaterials (sog. Memorisation)
- ▶ Verallgemeinerte Anwendung des abstrahiert erlernten Trainingsmaterials mithilfe von Wahrscheinlichkeitsberechnungen
- ▶ Reproduktion ggf. bei kurzen Snippets und detaillierten Prompts
- ▶ Aktuell: Vorlagefrage an EuGH, Rs. C-250/25, *Like Company v. Google*

4 Rechtliche Herausforderungen im Zusammenspiel von Open Source und KI

Output nach KI-Training mit FOSS

Urheberrechtsverletzung durch **Bearbeitung** im Output?

- ▶ § 69c UrhG nicht anwendbar, da kein menschlicher Bearbeiter
- ▶ § 23 Abs. 1 S. 1 UrhG nicht anwendbar, da keine Werksqualität des Outputs

4 Rechtliche Herausforderungen im Zusammenspiel von Open Source und KI

Output nach KI-Training mit FOSS

Auswirkung des Copyleft-Effekts auf KI-generierten Code

Copyleft = Software, die Bestandteile von FOSS-Programmkomponenten enthält oder von ihnen abgeleitet ist (= Derivative Work), darf nur unter derselben Lizenz weitergegeben werden. Der Effekt wird i.d.R. bei Bearbeitung oder Veränderung ausgelöst.

Bsp.: GPL-2.0, MPL-2.0

P: Derivative Work bei KI

insb. wenn kein copy paste Verfahren besteht (code completion bei Github Copilot)?

4 Rechtliche Herausforderungen im Zusammenspiel von Open Source und KI

Output nach KI-Training mit FOSS

Auswirkung des Copyleft-Effekts auf KI-generierten Code

Wann liegt ein Derivative Work bei KI vor, insb. wenn kein copy paste Verfahren besteht (code completion bei Github Copilot)?

- ▶ Mangelnde Rückverfolgbarkeit des KI-generierten Codes
- ▶ Einordnung kann abhängen von
 - ▶ Anteil des Codes, der unter Copyleft-Lizenz steht
 - ▶ Allgemeingültigkeit des Codes

→ Verantwortung für Compliance liegt bei Benutzer



5 Ausblick

→ Weiterhin rechtliche Herausforderungen



Mögliches technisches Vorgehen:

- ▶ offene Trainingsdaten (z.B. Common Corpus)
- ▶ Code unter permissiven Lizenzen verwenden

Mögliche rechtliche Änderungen:

- ▶ neue bzw. umfassende Regulierung, insbesondere Verlagerung der Verantwortlichkeit für Compliance
- ▶ dabei möglichst Ausgleich zwischen Innovationsförderung und Regulierung schaffen

Vielen Dank!



Melissa Irtel
Rechtsanwältin
JUN Legal GmbH

mirtel@jun.legal

