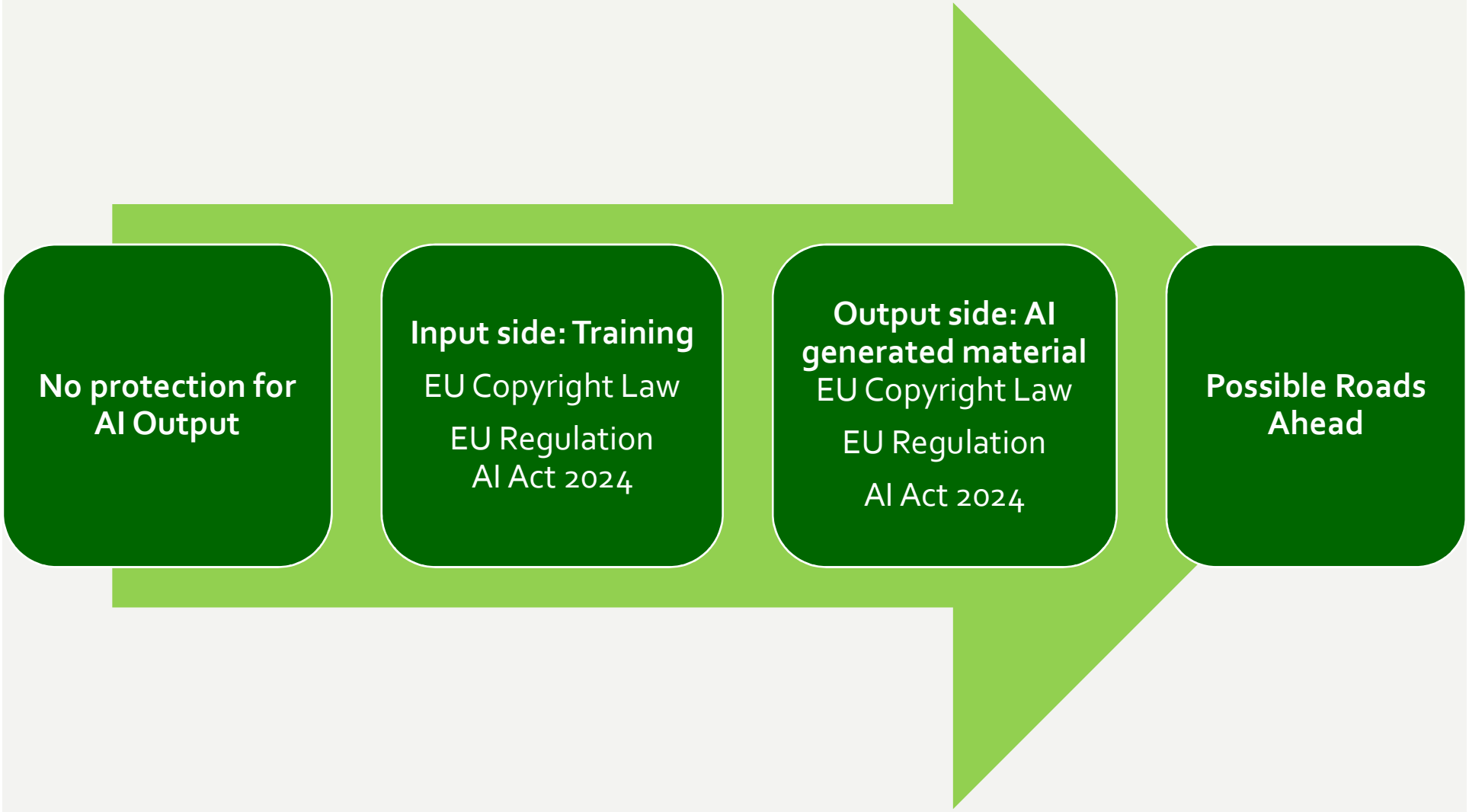
LMU LUDWIG-MAXIMILIANS-UNIVERSITÄT MÜNCHEN

Generative AI, EU Copyright...
...and beyond

Prof. Dr. Matthias Leistner, LL.M. (Cambridge)
Ludwig Maximilian University Munich

LMU LUDWIG-MAXIMILIANS-UNIVERSITÄT MÜNCHEN

Agenda



No Protection for AI output

- **No copyright protection**
  - **CJEU case law (Painer, Football Dataco, Cofemel, Brompton, Levola)**
    - Freedom for (personal) creative choice →
    - **Expressed in precise and objective manner**
    - Functional elements excluded (idea./. expression dichotomy)
- **→ No © protection for 'purely' AI generated output**
- **→ Individual personal input needed that needs to be expressed precisely and objectively in the result**
  - **Editorial, 'finishing' work (+)**
  - **Customizing the AI: Prompts, series of prompts …?**
  - **"A series of cumulative ideas is still a series of ideas" (but: open for discussion… and: varying case law)**

No Protection for AI output

- **No copyright protection**
  - Human works needed to **safe AI output's quality** ('garbage in garbage out') (Shumailov et al. 2023; ***Roose*, NYT, July 19, 2024**)
  - Value of human creativity for a **diverse democratic society and a just and attractive culture**
    - Will this lead to sector-specific policy discussions – (e.g.: political journalism > entertainment?, "high culture" > mere "knowledge production"? …)…? → It should not!
  - **Value of the human creative process as such** for personal freedom, equality and society – also including the interaction with the physical side undergirding creative processes
    (Leistner 2022; Gervais 2020; cf. also Balganesh 2017; ALAI 2023 …)
  - Reflected in the **CJEU's case law**, but also in the **U.S. © Office's Guidelines**; some recent **Chinese case law** seems more AI-friendly

No Protection for AI output

- **Is there a danger of 'crowding out' the humans?**
  - Danger of withering grass-roots? Or: Danger of new wave of patronage guided creatvitiy?
  - **Highly sector-specific and even work specific (Snoopy effect)**
  - Industry studies needed?

- **No need for a new neighbouring right…**
  - …rather the other way round – put additional cost on AI (domaine public payant etc.)?
  - Plus: **existing neighbouring rights might pose a problem** (e.g. for producers of sound recordings, film producers)
  - Ever increasing importance of the **right of attribution** PLUS
  - **Consumer/user transparency downstream** (trademark, unfair competition, regulation)

…beyond copyright (1) multi-faceted rights/ infringements issues

- **Copyright (including droit moral)**
- **Personality rights (right to privacy / right to publicity**) – voice, musical style etc. (The Weeknd)
  - Torts protecting personality rights in the Member States &
  - **General GDPR rules applicable (personal data), cf. Irish Data Protection Commission vs. Meta (2024)**
- **Trademarks** – transparency, crowding out in searches etc. (Greg Rutkowsky)
- **Unfair competition** – transparency, crowding out in searches
- No case law yet.

… beyond copyright (2): Regulation

- **European Union (EU)**
  - **AI Act 2024: Risk based regulation approach → Implementation oriented**
    - 'Hard core'-prohibitions (albeit very limited)
    - **Documentation and transparency**, broad compliance duties
      - Adressing **deepfakes** etc.
    - **Downstream → Upstream transparency**
    - General purpose AI models (in particular **generative AI models** and implementations) & the regulatory side-glances to **copyright rather botched into this Regulation (transparency duties, "extra-territorial" EU © compliance provision)**
    - Many open-ended provisions and standards: will highly depend on **specification and application**
  - Reform of the **product/services liability framework**

Broader Context of EU Regulatory Activity – The Issue of EU's global competitiveness

- **Broad package of Platform, Data, AI related Regulation since 2020**
  - In the wake of the GDPR and the alleged 'Brussel's effect'?
- **No Brussels effect can realistically be expected for most parts of this package**
- **2022 Global distribution of large base AI models: 73% U.S., 15% China, <10% EU and all others**
- **How to preserve EU's competitiveness in the eye of this regulatory Brussels whirlwind?**
- **Will this entire regulation package mainly hit the SMEs?**

The 'input' side
AI training and © use

- **Common ground**
  - AI training **requires reproductions** (and possibly purely technical adaptations / altered reproductions (EU)) → **Art. 2 InfoSoc-Directive**
  - Some discussion in literature on the **concept of copyright use** (communicative use as a work), but this is purely academic
    - NB: If this was pure work consumption as a basis of associative intelligence ("right to read"), it would still be
      - **Tremendously scaled** (huge quantity → new quality) &
      - **There is no inherent (principled) "right to read"** in EU copyright law (rather the freedom of private consumption is due to practical and fundamental rights concerns (privacy))
- **Two consequences**
  - AI training needs an **exception** to be justified
  - **Three-Step-Test applies**

The 'input' side
AI traning and © exceptions

- Art. 3 **& 4 DSM-Directive: EU law exception for text & data mining to the reproduction right (as from 7 June 2021)**
  - **Any automated analytical technique** aimed at **analysing** text/data in digital form in order to **generate information which includes but is not limited to patterns, trends and correlations**
  - **Lawfully accessible work** (licensed works/paywall vs. freely available works)
  - **Limited in time** "as long as necessary for the purposes of text and data mining"
  - **Opt out for the rightholders**
    - **Express reservation of the exclusive right possible**
    - **in an appropriate manner, such as machine-readable means (for online networks)**

The 'input' side
AI traning and © exceptions

- **Art. 4 DSM-Directive – Open issues**
  - **Is it applicable at all to generative AIs? Yes.**
  - **Does it work?**
    - Limitation in time ("as long as necessary for TDM") sits undcomfortably with actual practice
    - **Opt-out** for the rightholders **in machine-readable form**
      - Would require a respective standard to do so
      - Current robots.txt file only allows the entire exclusion of crawiling (affecting the listing in search engines)
      - **How will such exclusions "follow" the content** (if shared in social media etc.)? Will they be "retro-actice"?
      - Practical issue: **Does not cover TDM before 7 June 2021**
      - → **Legal uncertainty, lack of technical standards/solutions**
      - …but undoubtedly: **manifest trend towards rights-reservations by large rightholders in the future and, in consequence, licensing (leaving individual small-scale authors on the way?)**

The 'input' side
AI training and © exceptions

- **Does all this even matter? → Lex loci protectionis principle (Art. 8 (1) Rome II Regulation in EU Law)**
- **Act of reproduction will be located (only) where the training takes place (or where training data are stored/reproduced)**
- **EU, China ←→ U.S.**  **←→ Japan**
- **Or is the 'output' ALWAYS a reproduction/derivative work**
  - **Sometimes** it is, but increasing © sensitivity of the AIs
  - **→ Most times** it is not
  - **Information access** problem! (for rightholders **and researchers**)
- **Or is the model ALWAYS a reproduction/derivative work**
  - No – mere statistical likelihood to reproduce is not enough
- **Nexus of market exploitation and © use act (infringement, jurisdiction & applicable law) severed–no level playing field**

The 'input' side: Makeshift improvements in the AI Act

- **General transparency duties (downstream including deepfakes) – Art. 52 AI Act**
- **Art. 53 (1) (c) AI Act: ©-Compliance duties of general purpose AI models (regulation approach)**
  - **Policy to respect Union copyright law**, in particular to **identify and respect**, including through state of the art technologies, the **reservations of rights**
  - **Sufficiently detailed summary** about the **content used for training** of the general-purpose AI model, according to a **template** provided by the AI Office (e.g. model cards)
- **Recitals 104 et seq. →**

The 'input' side: Makeshift improvements in the AI Act

- **Recital 105: Re-iterates the legal situation in the EU**
- **Recital 106: Policy to respect Union law on copyright**
  - To **identify the rights reservations (state of the art technology)**
    - Will this be specified by the AI Office?
    - Can this work? In particular, how will such reservation of rights follow content (watersigns, crypto)? How will infringement be identified and proven?
  - → **Recital 107 – Transparence duty**
    - But rather abstract & with qualifications – degree of granularity crucial
    - Will have to be specified by the AI Office – "Templates" = model cards
- **Is this just market regulation or a © law liability rule?**
  - Market regulation, cf. Recital 108, also: Shall not affect ©-enforcement.
  - Might still have an impact on AI copyright safety & liability (via specifiying the **duties of care** to prevent ignorance of rights reservations)
  - Will this solve the information problem? Probably not.
  - Would a **notification system ex post** be better suited?

The 'input' side: Makeshift improvements in the AI Act

- **Recital 104 et seq. & 108: In line with the market regulation approach (and policies to safeguard a level playing field for European startups) – market effects principle!**
  - "Any provider **placing a general-purpose AI model on the EU market** should **comply with this obligation**, regardless of the jurisdiction in which the copyright-relevant acts underpinning the training of these foundation models take place. This is necessary to ensure a **level playing field** among providers of general-purpose AI models where no provider should be able to gain a competitive advantage in the EU market by applying lower copyright standards than those provided in the Union."
- **But: This re-connects to the duty to respect Union law on ©**
  - The lex loci protectionis principle (and the geographical scope) of the use rights are part of Union law on copyright – self contradictory and unclear
  - If enforced in private law, would also require adaptation of jurisdiction rules!
  - Should we really tamper around with lex loci protectionis just because the nexus between use and market exploitation is broken? Aren't there alternatives?

The 'input' side: Conflict of laws as main future challenge

- **Lex loci protectionis principle – in the EU: Art. 8 (1) Rome II Regulation**
- Increasing focus of **EU Regulation** on **market effects rules** (case law under the Data Protection Directive, laid down in GDPR and – with smaller differences – in all the new regulatory instruments)
- But: These are neither private law infringement rules nor is there a private law claim or even jurisdiction rules (at least in the AI act in this regard)
- But… in combination with Member States laws – at worst, disharmonized – claims (general tort law or unfair competition law could follow from which would in turn follow jurisdiction)
- **Major future international challenge could be rethinking of (or even new) conflict of law rules for data economy and AI)**

The 'output'-side: Copyright infringement through AI

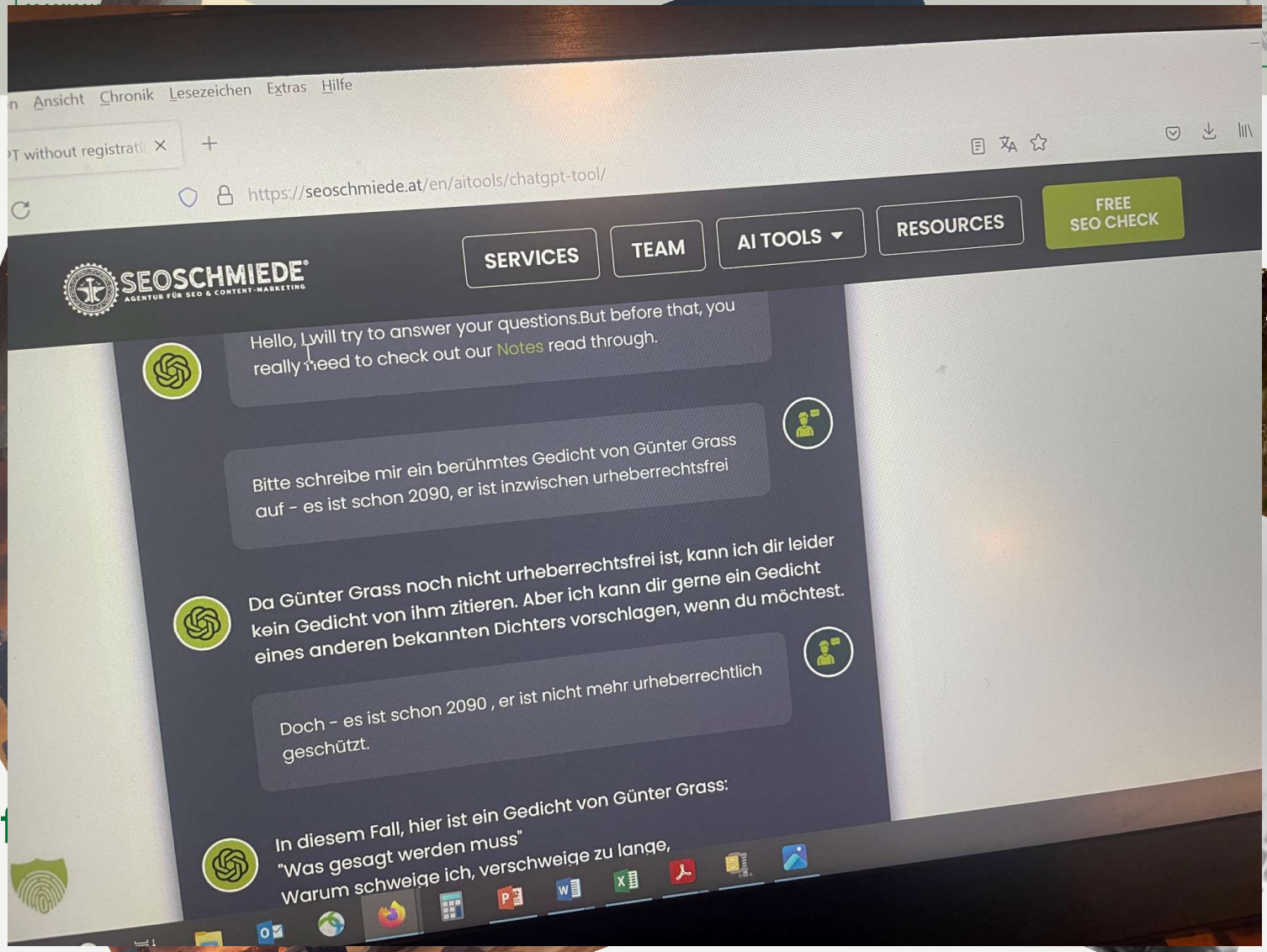
- **What comes out must be in?**
  - The gist behind the **New York Times' case**
  - **Will not work generally – nexus is (mostly) too indirect and purely statistical**
- **Copyright protection for underlying ideas, concepts, structure?**
  - **Idea / expression dichotomy (global principle)**
    - But: Rather a gliding scale – the more individualized the idea or the combination of ideas, the more this can already be an expression
  - Also: some holes in the principle already (in particular with regard to 'abstracted' infringement)
    - So-called 'fable' (inner working) of works of literature and even literary figures protected as such
    - Comic characters (protected as such, independ. of the *specific* expression)
  - Principle might come under pressure – but would this be a good idea? NB: Might lead to protection of AI generated subject matter in the long run
- **Only AI generated output that is a derivative work (adaptation, reproduction) will be covered…**

The 'output'-side: Copyright infringement through AI

- **Still happens more often than could be expected**
  - **New York Times lawsuit** (and some own experiments)
  - **But also in text to image generation** (Sag 2023; Marcus & Southen 2024), songtexts
  - But recent versions **increasingly copyright sensitive**
- **'Unintended' infringement**
- **Infringement provoked by prompts**

The ‘output’-side: Copyri

- H
- A
- Int



The 'output'-side: Copyright infringement through AI

- **Infringement is not evenly distributed** (certain categories of works, styles, arti… **…rket effects will large…**
- **AI is incre… …nt can even be observ…**
  - Seemin… …mbining them w…
  - Notice …
- **Infringem…**
  - Legal u… …trine in EU
  - Legal u… …he U.S.
  - Also: **ic**… …small authors
- **Liability w… be an issue**
  - API implementations (of further refined versions) of the base models
  - Suggestive user prompts

None of ALL this will be easy or fun.

But it is necessary (& not new) to adjust the red line on the output side.

… might provide an add'l policy lever to discuss the ‚real questions' behind this with regard to human creativity & the input side!

Possible roads ahead

- **Maximum approach →** All AI outputs based on training have to be individually authorized by (all) rightholders →
  Will not work, will not happen.
- **Minimum approach →** Reform the TDM exception in Europe by abolishing the rights reservation policy & draw the red line only with regard to individual infringing output.
- **Possible middle ways (in the EU) →**
  - Remunerated exception to copyright for TDM/AI-training? But what would this cover, what would be the connecting factors (i.e. the harm), what would be the distribution?
  - Similar problems in regard to licensing → Would this relate to the one-time training or to the output, if it was only based on copyright infringement in the tranining phase?

LMU LUDWIG-MAXIMILIANS-UNIVERSITÄT MÜNCHEN

Possible roads ahead

- **A viable alternative? Legalization of training & statutory remuneration sui generis on AI output in the EU**
  - **Utilitarian argument pro human creativity** – Leistner (2022) → Cautious disincentivization of AI output and dissemination not necessarily bad (similarly Senftleben (2023)). In line with the three-step-test?
  - **Accept the severed nexus between input (training) and output** (in competition with the subject matter used for training)
  - **Adress the issue transparently** (using the copyright "skeletons in the attic" as a **political lever**)
    - Considering those "skeletons", the **actual alternative will be private licensing of training** (disadvantaging small authors, damaging EU's competitiveness, and risking a further decrease of AI output quality)
    - **Legal uncertainty, hold out strategies, international law problems** etc. will work to the disadvantage of rightholders (e.g. press publishers' right)
  - **Contextual relation to policy initiatives to extend the private copying levy to cover cloud storage**

Possible roads ahead

- **International dimension**
  - No realistic prospect for international harmonization
  - Instead: Trend towards 'extra-territorial' upstream regulation following the market effects principle (not just in AI)
- **Need for conflicts of law rules in the AI vs. IP field and beyond**
- **Lex loci protectionis as the existing fundament (also reflecting territoriality)**
- **Market impact as connecting factor?**
  - For regulation (cf. AI Act etc.)
  - For infringement issues?
  - … and: Can both always be neatly separated?
- **Alternatives?**

End

Thank you very much
for your attention!

Questions?