



# RECHT, KOGNITION & DER AUFSTIEG KÜNSTLICHER INTELLIGENZ

*Michael Epping*      *Christoph Winter*

Die Bewohner des römischen Reiches hatten keine Probleme damit, zur Erheiterung der Massen Menschen den Löwen im Kolosseum zum Verzehr vorzuwerfen. Die Verfassung der Vereinigten Staaten aus dem Jahr 1787 legitimierte die Sklaverei und in einigen hinduistischen Gemeinschaften in Indien, Nepal oder auf Bali war es üblich, Witwen gemeinsam mit der Leiche ihres Ehemannes zu verbrennen (sog. *Sati*). Doch so selbstverständlich wie diese jeweiligen Handlungen den Menschen der Zeit erschienen, so selbstverständlich widersprechen sie den Intuitionen der Mehrzahl der Menschen heute.

Da sich für jede Epoche und Region eine Vielzahl an grausamen Handlungen finden und kritisieren lässt, die im Einklang mit dem allgemeinen Zeitgeist standen, fragt sich, welche derzeit vorherrschenden Intuitionen in absehbarer Zeit mit der Selbstverständlichkeit als moralisch verwerflich bewertet werden, mit jener die Menschen heute Sklaverei und Witwenverbrennungen verachten. Kann etwa die globale Armut, wie *Pogge* immer wieder betont, tatsächlich als Verbrechen gegen die Menschlichkeit eingestuft werden? Kann noch immer an der Kriminalisierung von Inzest (§ 173 StGB), Cannabiserwerb (§§ 29 ff. BTMG) oder der Nichtkriminalisierung der Massentierhaltung festgehalten werden? Inwiefern schützt und sollte das Strafrecht gesellschaftliche „Tabus“ schützen (vgl. *Hörnle* 2005)? Können und sollten moralische sowie juristische Bewertungen überhaupt übereinstimmen?

Diesen und ähnlichen Fragen zum Rechtsintuitionismus nähert sich das Seminar auf Grundlage der sog. *dual-process theory*, welche spätestens durch *Daniel Kahnemans* „Thinking fast and slow“ (2012) auch über die Entscheidungswissenschaften hinaus an Einfluss gewann (für einen Kurzüberblick siehe [hier](#); ein Kommentar von *Kahneman* findet sich [hier](#)). Zentrale These ist die Unterscheidung zwischen zwei Arten des Denkens: das schnelle, instinktive und emotionale „System 1“ und das langsamere, rationalere und logischere „System 2“. Während das überwiegend angewandte schnelle Denken sehr effektiv ist, kann es unabhängig der Intelligenz oder Expertise zu gravierenden *systematischen*

Denkfehlern führen. Ziel ist über die psychologische Analyse des menschlichen Denkens zu ergründen, welche Gedanken vertrauenswürdig sind und welchen Intuitionen wir besonders skeptisch gegenüberstehen sollten.

Dass die Entscheidungen des Menschen nicht immer rational sind, sondern auch auf vorhersehbaren Denkfehlern basieren, hat bereits zu Forschungsergebnissen mit erheblichen Implikationen für Rechtsanalyse, -Verständnis und -Rechtfertigung geführt. Wir möchten uns diesem auf dreifache Weise nähern. Neben den bereits erwähnten und grundlegenden strafrechtsphilosophischen Aufgabenstellungen, welche die Legitimation von Strafnormen untersuchen, stellt sich die Frage, ob jene Denkfehler instrumentalisiert werden können und sollten, um Gutes zu bewirken. Das vom letztjährigen Nobelpreisträger Richard Thaler und Cass Sunstein geprägte *Nudging*, „eine Methode, das Verhalten von Menschen auf vorhersagbare Weise zu beeinflussen, ohne dabei auf Verbote und Gebote zurückgreifen oder ökonomische Anreize verändern zu müssen“ (*Thaler/Sunstein* 2008), bietet einen vieldiskutierten Ansatz. Aber ist eine solche Beeinflussung verfassungsgemäß und welche (rechts)ethischen Faktoren spielen diesbezüglich eine Rolle?

Einige Philosophen, etwa *Peter Singer* (2014/2017) und *Joshua Greene* (2014), gelangen (auch) auf Basis der *dual-process theory* zu der Ansicht, dass menschlichen Intuitionen aufgrund ihrer Fehleranfälligkeit (weitestgehend) nicht getraut werden kann. Der dritte Abschnitt dieses Seminars möchte daher die Alternativen menschlicher Intuitionen ergründen. Neben möglichen Strategien zur Milderung von Vorurteilen (*debiasing*) werden dabei insbesondere die Vor- und Nachteile KI-basierter Entscheidungssysteme beleuchtet.

Inwiefern kann künstliche Intelligenz etwa den Prozess richterlicher Entscheidungsfindung verbessern? Schließlich haben von der Verhaltensökonomie inspirierte Untersuchungen gezeigt, dass die Urteile von RichterInnen noch immer von rechtlich irrelevanten Faktoren wie der Uhrzeit oder der Zahl eines Würfels beeinflusst werden (siehe dazu [hier](#); für einen Kurzüberblick siehe [hier](#) oder ausführlicher [hier](#)). Gleichzeitig wird vermehrt in den USA mit Hinweis auf *algorithmic biases* diskutiert, inwieweit Vorurteile durch die Anwendung künstlicher Intelligenz lediglich manifestiert werden. Können die Probleme der Diskriminierung durch Algorithmen langfristig gelöst werden? Welche (verfassungs-) rechtlichen Vorgaben gilt es, hierbei zu beachten? Würde ein KI-basiertes Rechtssystem den Zugang zu Recht und hiermit die Bedingungen des international anerkannten Prinzip des *fair trials* stärken, wie China (vgl. dazu [hier](#)) und Russland (vgl. dazu [hier](#)) in ihren jüngsten Initiativen argumentieren?

Welche Gefahren, Chancen und rechtliche Regulierungsbedürfnisse bringt die generelle Verschiebung der Autorität von menschlichen Intuitionen zu „künstlichen“ Algorithmen mit sich? Welche rechtsethischen und psychologischen Besonderheiten gibt es etwa bei der Regulierung autonomer Fahrzeuge und Waffensysteme zu bedenken?

Neben diesen Fragen der Gegenwart und nahen Zukunft wird ein letzter Abschnitt jene der fernen Zukunft behandeln. Welche rechtlichen Möglichkeiten bestehen etwa, um *digitale Diktaturen* zu verhindern? Ist dies überhaupt eine Gefahr? Gibt es rechtliche Wege um sicherzustellen, dass eine „Superintelligenz“ (Bostrom 2017) menschliche Werte vertritt? Wäre dies überhaupt wünschenswert? Sollte nicht nur künstliche Intelligenz, sondern auch ein künstliches Bewusstsein entwickelt werden, fragt sich, einerseits, ob es moralisch schutzwürdig ist und andererseits, welche rechtlichen Möglichkeiten bestehen. Können die Menschen diesbezüglich ihren Intuitionen vertrauen?

**Das Seminar bietet die Möglichkeit, über die klassische Rechtsdogmatik hinaus Einblicke in die Schnittstellen von (Moral-)Psychologie, praktischer Philosophie, Recht und künstlicher Intelligenz zu erhalten, wobei kein Vorwissen vorausgesetzt und entsprechende Hilfestellung gegeben wird. Aus interdisziplinärer Perspektive soll gemeinsam diskutiert werden, wie künstliche Algorithmen und menschliche Intuitionen die Gestaltung und Auslegung von Recht *beeinflussen* und *beeinflussen sollten*.**