



**Online Only**

RÜDIGER STIX

## „Neutrals in the Net“

oder:

wie intelligent ist KI/AI im Cyber-War?

### **Abstract**

Eine Denkschrift zu den konkret beforschten Prioritäts-Technologien und deren Rechtlichen Rahmenbedingungen im öffentlichen Diskurs und in der EU - Defence Research (Stand Ende Jänner 2018).

## Inhaltsverzeichnis

---

	Seite
<b>Abstract</b>	1
Inhaltsverzeichnis	2
<b>Kapitel 1</b>	3
<b>Einleitung: aktueller gesellschaftlicher Diskurs</b>	
<b>Kapitel 2</b>	10
<b>Zentrale Begriffe, Themenfelder und Fragen</b>	
<b>Kapitel 3</b>	19
<b>Wie „intelligent“, wie „einsichtig“ oder „weise“ ist Homo Sapiens? Und wie „intelligent“ ist KI/Künstliche Intelligenz?</b>	
<b>Beispiele aus der Forschung, Einsatz von Gehirnstromtechnik aktuell</b>	26
<b>Zusammenfassung</b>	26

## Kapitel 1

---

### Einleitung: aktueller gesellschaftlicher Diskurs

Künstliche Intelligenzen kennt jeder, zumindest in öffentlichen Diskussionen und in Fachdiskursen, und in den dramatischen Bildern zahlloser Science Fiction Filme. Sie haben unsere Vorstellungen geprägt - von der „*Space Odyssey*“ über das „*Sky Net*“ im „*Terminator*“ bis zu Internetwelten der „*Matrix*“, nicht zu vergessen für die jüngeren Generationen die interaktiven Welten der Simulationsspiele, meistens in sehr militärischen Szenarien.

Auch Mensch/Maschine Schnittstellen, einschließlich der BCI/Brain Computer Interfaces finden wir zwischen Cyborgs, „*Iron Man*“ und „*Avengers*“ oder „*Avatar*“.

Die gesellschaftlichen und auch die wissenschaftlichen Begleitdebatten schwanken dabei stets zwischen Beschwichtigungen und Alarmismus - also zwischen der beruhigenden Behauptung, dass es sich lediglich um Phantasie-volle „*Science Fiction*“ handelt, und den Weckrufen, dass wir es schon lange mit „*Science*“ zu tun haben, und nicht mehr mit bloßer „*Fiction*“.

Aktuell stehen wir aber unwiderruflich in den Grenzbereichen der Verteidigungsforschung auch in Europa, weltweit in der Medizin, und in den fehlenden Antworten aller unserer Rechtssysteme vor realen Aufgaben, mit denen wir jetzt mehrfach Neuland betreten haben:

So macht die Verteidigungsforschung traditionell in den USA, und nunmehr aktuell in der EU mit der Künstlichen Intelligenz (KI), der Digitalisierung im Cyberspace, der Robotik und Sensorik und den Neurokognitionswissenschaften die Mensch/Maschine Schnittstellen zu den wichtigsten Prioritäten ... und damit zum absehbar größten Budgetanteil der kommenden EU Forschungsförderung nach dem jetzt auslaufenden Rahmenprogramm „*Horizon 2020*“ ... und zu einem Milliardenmarkt.

Auch das „*Sky Net*“ Szenario einer unkontrollierbaren Netzintelligenz aus den „*Terminator*“ Filmen wird schon länger ernst genommen, nicht nur von Naturwissenschaftlern wie Stephen Hawking, oder auch von Ökonomen, die mit Yuval Noah Harari über „*Transhumanismus*“ diskutieren. Schließlich hat auch im Militär niemand geringerer als der zuständige General der US-Streitkräfte, Paul J. Selva, als Vice Chairman der Joint Chiefs of Staffs (<http://www.defense.gov/News/Article/Article/927792/dod-studies-terminator-weapons-conundrum-selva-says>) seit über einem Jahr und wiederholt öffentlich vor dem „*Terminator/Skynet Conundrum*“ gewarnt.

General Selva weist dabei ziemlich ungeschminkt auf zwei Entwicklungen hin:

Erstens ist beim derzeitigen Entwicklungsstand der möglichen Technologien ein unkontrolliertes, und vielleicht absehbar auch unkontrollierbares autonomes System im und aus dem Netz, in etwa nach dem populär-fiktionalen Bild eines „Skynet“ aus der Terminator-Filmreihe, sehr wohl real denkbar, wobei eine autonome künstliche Intelligenz im Netz nach dem Modell des „Skynet“ naturgemäß viel bedrohlicher ist - im Film wie in der Realität...

Zweitens sind die Streitkräfte generell, und auch das US Militär selbst, nicht in der Lage, eine derartige Entwicklung mit Sicherheit zu verhindern. Der Grund hierfür liegt schlicht und einfach in der Verteilung der Ressourcen:

Wenn man die klassischen strategischen Programme vom „Manhattan Project“ zur Entwicklung der Nuklearwaffen bis hin zu den entscheidenden Schritten in der Luft- und Raumfahrt in den vergangenen Jahrzehnten betrachtet, so waren öffentliche Dienststellen, also Behörden der US Regierung federführend – US Militär und staatliche Forschungsagenturen von der DARPA bis zur NASA.

Im Bereich der Technologie von Künstlicher Intelligenz, ist es – weltweit – genau umgekehrt: die Masse der Ressourcen liegt bei internationalen Konzernen wie Google oder Facebook.

Die US-Regierung und insbesondere das US-Militär können dabei eigentlich, so General Selva, nur eine Art „Triage-Funktion“ wahrnehmen, und darauf achten, dass Entwicklungen steuerbar bleiben.

Gleichzeitig macht gerade die Militärmedizin revolutionäre Fortschritte im Bereich von Mensch/Maschine-Schnittstellen, die laufend klinisch nutzbar werden, vor allem in der Behandlung traumatischer Gehirnschäden.

Dies reicht von einer neuroprothetisch gestützten Kommunikation durch „*silent talking*“ sogar mit „*Locked-In-Patienten*“, wie auch zu nicht-invasiven Interventionen in beide Richtungen: etwa in der persönlichkeits-modifizierenden Behandlung durch nichtinvasive Gehirnstromtechnik bei schweren posttraumatischen Persönlichkeitsbeeinträchtigungen.

Dies bedeutet, dass wir uns auch der zweiten neuen Herausforderung stellen müssen, neben den Programm-technischen IT – Fragen des Netzes den Fragen der Neurokognition: Wenn wir uns ab jetzt, zumindest in der Kommunikation mit Patienten oder mit Einsatzkräften, über BCI/Brain Computer Interfaces auf die Hilfe von hochwirksamen Künstlichen Intelligenzen abstützen, deren Wirkung wir kaum vorhersehen können, dann müssen wir laufend realistisch bewerten „wie intelligent ist KI“?

Damit beschäftigt sich dann der Hauptteil unserer Diskussion im dritten Kapitel an Hand der Beurteilung, in wie weit eine gesicherte Simulation von intelligenten Entscheidungen nicht nur unter strikt vorgegebenen Regeln, sondern auch von impliziten Wissen und Können möglich ist.

Gleichzeitig müssen wir zu den klassischen ethischen Dilemmata parallel ergänzend die medizin- und rechtsethisch neue Frage beantworten, wie lange denn „die Gedanken frei sind“, wenn sie etwa durch Brain-Computer-Interfaces abgegriffen werden können und müssen, bevor ein rechtlich-verbindlich zurechenbarer Ausdruck etwa des Patientenwillens auch nur halbwegs zweifelsfrei ist (wozu es umfangreiche Überlegungen gibt, in Österreich von insbesondere von *Likar und Pinter, ÖGARI, 2017*).

Gesellschaftlich stehen wir natürlich in einer noch wesentlich umfassenderen Debatte. Sie reicht aktuell „vom „Davos Man“ zum „Homo Deus“, und von Robot-pflichtigen Leibeigenen (in Österreich bis ins Jahr 1848) bis zur Herrschaft genetisch optimierter Cyborgs in künstlichen intelligenten Netzen ...“ (vgl. [scienceblog.sfu.ac.at](http://scienceblog.sfu.ac.at) vom 29.1.2018), und geht verständlicherweise weit hinaus über die aktuell brennenden Fragen des

- Neutralitätsrechtes,
- des Sanitätsrechtes und der
- Einsatzbestimmungen für Militär und FR/“First Responder“
- unter Einsatz der Prioritätstechnologien aus der Europäischen Verteidigungsforschung.

Zu diesen angeführten generellen globalen Auswirkungen diskutierte pünktlich zum Wintereinbruch in den Schweizer Westalpen im Jänner 2018, wie jedes Jahr, das Weltwirtschaftsforum in den Schweizer Alpendörfern Klosters und Davos, wobei Davos einen eigenen Typus Mensch charakterisiert: Der legendäre „Davos Man“ ist die politisch globalisierte Ausgabe der „Masters of the Universe“ aus der Wallstreet. Der „Davos Man“ wurde weltweit berühmt durch Timothy Garton Ash (2005, in „*The Guardian*“), und definiert etwas karikierend den durchschnittlichen Besucher des Weltwirtschaftsforum im wunderschönen Schweizer Alpendorf Davos, als einen älteren, männlichen Angelsachsen aus einer internationalen Finanz- und Politelite, in der sich natürlich, so spottet Ash, auch „als Davos-Minderheiten Inder, Chinesen, Afrikaner und/oder Frauen“ befinden. Timothy Garton Ash beruft sich dabei auf einen der Klassiker unter den Superstars der US–Geopolitik, den ehemaligen Direktor des Harvard Center for International Affairs, Samuel Huntington, Sicherheitsberater unter Jimmy Carter (Democrat, Präsident der USA von 1977 bis 1981), und dessen Skepsis, dass „der Davos-Man eine globale Elite repräsentiert, die kaum nationale Loyalitäten kennt, und die Staaten und ihre Regierungen nur dann als nützlich betrachtet, wenn sie Operationen der globalen Eliten umsetzen“.

Aber lassen wir angemessen staunend die politischen Debatten im heurigen Davos des Jahres 2018 vorerst offen, auch wenn erstmals mit Donald Trump ein US-Präsident aufgetreten ist. Dementsprechend richteten sich die Augen der Welt auch eher auf die Wirtschaftsbeziehungen und die globalen Konflikte der Regierungen zwischen den USA und China, obwohl sich auch einer von Donald Trumps vehementesten innenpolitischen Gegnern und wohl weltweit umstrittensten Multimilliardären, George Soros, zum drohenden Untergang von Kulturen geäußert hat, und in bemerkenswerter Weise vor allem den US-amerikanischen IT-Konzernen wie Google und Facebook explizit vorwirft, dass sie durch ihre Kooperation mit „totalitären Regimen“ (wobei er Russland und China ausdrücklich nennt) ein Netz von totalitärer Kontrolle schaffen würden, ein *“web of totalitarian control“*.

In Großbritannien wird aus genau diesem Grund von der Regierung im Rahmen des Nationalen Sicherheitsrats eine eigene Struktur eingerichtet, die *„Fake News und Desinformation“* bekämpfen soll. Diese Kommunikations-Organisationseinheit soll die „nationale Sicherheit“ gegen staatliche Desinformanten gegnerischer Mächte und andere nichtstaatlicher Akteure verteidigen, was im Rahmen der Diskussionen über das Verteidigungsbudget in London bekanntgegeben wurde. Es dürfte sich also tatsächlich um eine militärisch überwachte Struktur handeln, die in Zukunft auch den zivilen Kommunikationsverkehr überwachen und regeln soll.

Theresa May selbst benutzte daher auch Davos, um ihren Kampf gegen Soziale Netzwerke fortzusetzen, den sie von ihrem Vorgänger David Cameron übernommen hatte. In Davos ging es Theresa May offenbar auch darum, die dort versammelten Großinvestoren von der britischen Position zu überzeugen, denn in einer Runde von Industrievertretern, die von US-Präsident Donald Trump zum Dinner eingeladen worden waren, waren mit Siemens und SAP nur deutsche Technologiefirmen der klassischen Industriesektoren, aber kein US-Internetkonzern präsent.

Diesmal war neben den *„Davos Men“* und den *„Davos Women“* aber auch ein anderer, sehr hartnäckiger Gast erstmals in Davos: *„Homo Deus“*, präsentiert von seinem Schöpfer *Yuval Noah Harari*, dem seit zwei Jahren unbestreitbar weltweit am meisten gelesenen Militärhistoriker. Ob ein *„Homo Deus“* der Künstlichen Intelligenzen den Menschen verdrängt oder ihn komplett ersetzt, ist keine ganz neue Frage. In Davos standen daher auch der grundsätzliche Umgang mit den Künstlichen Intelligenzen und mit Big Data zur Diskussion, und vor allem die Auswirkungen auf die Arbeitswelt und unsere Gesellschaften in den kommenden Jahren.

Wie es sich für ein „Weltwirtschaftsforum“ geziemt, gibt es dazu auch eine aktuelle Resolution, die wohl dem Anlass entsprechend sehr allgemein gehalten, aber fairerweise angemerkt, ebenso allgemein akzeptabel ist:

*“As Artificial Intelligence (AI) and the Internet of Things (IoT) coalesce to create unique outcomes and opportunities, the need to strengthen the trust, transparency and fairness of these socio-technical systems is central for a balanced Fourth Industrial Revolution.*

*The challenges are multidimensional. Artificial Intelligence systems (and the platforms and IoT sensors which feed data into them) are impacting governments, enterprises, individuals and institutions in multiple ways. From transportation, to health, to finance, to commerce, the effects of these systems are felt but rarely seen. Whether by accident or design, these systems operate in the background of our lives. Their lack legibility and transparency creates a fog of understanding over how AI and IoT are shaping society. Our ability as a society to recognize, understand and correct harmful AI and IoT societal impacts is limited. Progress is stalled.*

*A more holistic set of governance frameworks for AI and IoT systems is needed. The overwhelming complexity, concentrated power asymmetries, competing incentives and global data flows create a new set of challenges which our existing institutions were not designed to address. A shared taxonomy of risks, pragmatic impact assessment tools and enabling policy frameworks are needed to ensure the promise of the Fourth Industrial Revolution can be realized” (Protocol Design Networks on AI, IoT and the Future of Trust; this project is part of the World Economic Forum’s Shaping the Future of Digital Economy and Society System Initiative).*

Deutlich konkreter wurde jedoch C. Vijayakumar, CEO von HCL, einem multinationalen Anbieter von IT-Dienstleistungen mit Sitz im indischen Noida. Sein Konzern beschäftigt ca. 120.000 Menschen, und hat größte Probleme, weitere ausgebildete Mitarbeiter zu finden: “In jedem Technologieunternehmen mangelt es an qualifizierten Fachleuten”, erinnert Vijayakumar, und vermutet rund eine Million Jobs im IT-Sektor, für die es einfach nicht die richtigen Fachkräfte gibt.

Auf der anderen Seite stehen Millionen Menschen, auch in den entwickelten Industrienationen der USA und Europas, die keinen angemessenen Job mehr finden. Einer der häufigsten Gründe ist dabei natürlich, dass sie nicht die Fähigkeiten mitbringen, die in einer globalen und digitalen Welt von Nöten sind - war man sich in Davos einig.

*“Früher wurden die Arbeiter ausgebeutet”* erinnerte Yuval Noah Harari, als Historiker an der Hebräischen Universität von Jerusalem, und Schöpfer des Weltbestsellers „*HomoDeus*“ sowie Autor von Bestsellern wie *“Eine kurze Geschichte der Menschheit”*. *“Jetzt fürchten die Menschen, dass sie völlig irrelevant zu werden”,* so Yuval Noah Harari, denn wenn ich ausgebeutet werde, dann bin ich wenigstens irgendwie wichtig: *„Sie können mich erschießen, aber wer macht dann meinen Job? Wenn ich völlig irrelevant bin, dann ist das viel Angst erregender.“*

Yuval Noah Harari macht auch kein Hehl daraus, dass er zwar von seinen Analysen überzeugt ist, aber keine Antworten anbieten kann. Auch gesicherte Prognosen vermeidet er: „Haben wir noch zwei Jahre Zeit? Zwanzig Jahre oder Zweihundert Jahre?“ ... war eine konkrete Frage. Yuval Noah Harari war sich aber in seiner Antwort in zweierlei Hinsicht ziemlich sicher: In zwei Jahren, so meinte er, wird sich die Welt, so wie wir sie kennen nicht allzu sehr geändert haben ... *„in zweihundert Jahren wird die Welt aber derart geändert sein, dass die dann*

*herrschenden Spezies sich von uns etwa in dem Ausmaß unterscheiden, wie wir von den Neandertalern ... oder den Schimpansen“...*

Yuval Noah Harari lässt keinen Zweifel daran, dass alles für eine überwältigende Dominanz künstlich intelligenter Systeme spricht, im Guten wie im Bösen, aber er hütet sich davor, der Menschheitsgeschichte ein eschatologisches Ziel zu unterstellen: *„Von Kanus sind wir erst auf Galeeren, dann auf Dampfschiffe und schließlich auf Raumschiffe umgestiegen, doch wir wissen immer noch nicht, wohin die Reise gehen soll. Wir haben größere Macht als je zuvor, aber wir haben noch immer keine Ahnung, was wir mit ihr anfangen sollen.“*

Auch hierzu ist Yuval Noah Harari als Militär- und als Universalhistoriker eher nüchtern. In all den Jahrtausenden, in denen Landbesitz entscheidend war, so Yuval Noah Harari, hat sich eine Art von adeliger Grundherrschaft herausgebildet, und in den letzten zweihundert Jahren der Industrialisierung eine Oligarchenklasse der „Kapitalisten“. Der unüberbrückbare Unterschied aber in den nächsten hundert, zweihundert Jahren, so Yuval Noah Harari, wird darin bestehen, dass *„die genetisch weiterentwickelten und mit künstlicher Intelligenz verstärkten neuen herrschenden Klassen eben keine Klassen von Menschen über Menschen sind, sondern neue Spezies...“*

Keines unserer Rechtssysteme, und keine unserer gesellschaftlichen Institutionen ist darauf eingerichtet...und schon gar nicht auf die sich derzeit aufbauenden asymmetrischen Konflikte - ganz zu schweigen von echten kriegerischen Auseinandersetzungen unter dem Einsatz von Ergebnissen aus jenen Prioritätsfeldern, welche Mensch und Netz verschmelzen lassen.

Immerhin gedenken wir heuer als Jahresregenten nicht nur der ersten hundert Jahre, in der sich das alte Europa auf dem Höhepunkt weltweit führender Wissenschaft und wohl auch maßgeblicher zeitgenössischer Kultur torkelnd und „Schlafwandelnd“, um Christopher Clark zu zitieren, in die Hölle und in die Mutterkatastrophe des Ersten Weltkrieges gestürzt hatte, mit allen darauf folgenden Kaskaden des Grauens.

Wir gedenken heuer auch des Jahres 1848, in dem mit der bürgerlichen Revolution auch endgültig die Leibeigenschaft aufgehoben worden ist und damit die Pflicht zum „Robot“, zu „Hand- und Spanndiensten“ für die Grundherrschaft, deren Ländereien man nicht verlassen durfte: eine der ersten und großen Erfolge der Revolution von 1848/9, eingeleitet durch Hans Kudlich, und aufbauend auf die Abschaffung der Leibeigenschaft, mit der schon Josef II., ältester Sohn Maria Theresias nach seinem Amtsantritt als römisch deutscher Kaiser begonnen hatte. Dieser schrieb auch Geschichte durch das Toleranzpatent, mit dem erstmals auch Protestanten und Juden wieder zu einer höheren Schulbildung zugelassen worden sind. Ein weltweit herausragendes Exempel der Rechtssetzung hat dabei ebenfalls seine Wurzeln in der Theresianisch/Josephinisch/Leopoldinischen Reformperiode: das Westgalizische Gesetzbuch, so wie wir es heute noch im Kern als unser bürgerliches Recht im geltenden ABGB pflegen. Dieses Westgalizische Gesetzbuch wurde von erfahrenen naturrechtlich aufklärerisch motivierten Juristen unter der Leitung von Karl Anton von Martini meisterlich konzipiert und experimentell im Jahre 1797 in Westgalizien in Kraft gesetzt.



Nach dieser erfolgreichen Erprobung unter realen Bedingungen in Westgalizien, und klug - um nicht zu sagen weise - ergänzt zum ABGB durch Franz von Zeiller, wurde es endgültig im Jahr 1812 für die deutschen Erbländer der Österreichischen Monarchie mit Patent in Kraft gesetzt, noch während der Napoleonischen Kriege, diesmal durch den ersten österreichischen Kaiser, Franz I., der als letzter römisch – deutscher Kaiser, als Franz II., den Reichsdeputations-Hauptschluss durchführte, um Napoleon den Griff auf die römisch–deutsche Kaiserkrone zu verwehren.

Fairerweise müssen wir bei allem Stolz über die österreichischen Reformer in der Schaffung naturrechtlicher Rechtssysteme anmerken, dass sowohl Napoleon, oder das dann im deutschen Bund gegen Österreich konkurrierende Preußen nicht weit zurücklagen: der Französische Code Civil unter Napoleon trat 1804 in Kraft, und die großen Reformer Preußens, Karl Freiherr vom und zum Stein, sowie Karl August von Hardenberg konnten sich mit ihren umfassenden Reformen durchsetzen nach der vernichtenden Niederlage Preußens gegen Napoleon bei Jena und Auerstädt.

Fügt man zu dieser Aufzählung bedeutender europäischer Aufklärung als Basis unserer heute bestehenden kontinentalen Rechtssysteme noch Alexander von Humboldt noch beispielhaft für seine umfassende Konzeption einer „Humboldtschen Universität“ mit dem Anspruch der „universitas“ hinzu, oder auch Carl von Clausewitz, als den bedeutendsten strategischen Denker der Neuzeit, so erhält man ein recht gutes repräsentatives Beispiel für die Ideen und Konzeptionen, die unsere heutigen Institutionen und Rechtssysteme ausmachen, wobei die angelsächsischen Entwicklungen natürlich ebenfalls mit einbezogen werden müssen.

Wir befinden uns also in der Epoche, in der unser heute geltendes Recht und die wesentlichen Institutionen geschaffen wurden, die vor allem mit den ersten zwei Problemen von Yuval Noah Harari umgehen mussten: mit den alten Grundherrschaften und den neuen Industrien, mit Geburtsadel und freien Bürgern, und mit Kohle und Stahl, mit Öl und mit elektrischem Strom auf Schienen und auf Schiffen, mit Wehrpflichtigen und mit den Arbeitern in den Fabriken und den Kraftwerken.

All diesen in dieser Epoche entwickelten und Respekt einflößenden - und bis heute dominierenden - Gedankengebäuden, und allen darauf bis heute aufbauenden Institutionen, von unseren Parlamenten bis zu unseren Gerichten, von den frei ausgeübten Gewerben, der Industrie und dem Handel, in frei gebildeten Vereinen, Parteien oder Gewerkschaften, in Akademien der Wissenschaft, den Kliniken, den Stäben und Generalstäben oder in den Redaktionen vom Fachjournal bis jetzt zum Internet, hat damals niemand im Traum daran gedacht, wie wir mit künstlichen Intelligenzen der unterschiedlichen Entwicklungsstufen umgehen, und wo wir uns Menschen abgrenzen gegenüber den Geräten und den Netzen, mit denen wir uns mit Hilfe unseres Gehirnstromes verbinden.

Es ist dennoch bemerkenswert, dass nach Europas selbst verschuldeten Massengräbern des Ersten Weltkrieges vor hundert Jahren unsere europäischen Kulturen derart zerreißen konnten, dass Bolschewismus und Faschismus sich in allen ihren totalitären Formen durchsetzten, in historischer Reihenfolge als Leninistisch-Stalinistisch real existierender Sozialismus, als Faschismus und als Nationalsozialismus Hitler´scher Prägung, und dann

Leninistischer real existierender Sozialismus Mao's, und nach all und inmitten von den –zig Millionen an Toten und Vertriebenen, die Schreckensvisionen und die Dystopien eines George Orwell oder eines Aldous Huxley's auch dadurch an Beachtung gewannen, als zunehmend Technologien der Überwachung oder der mentalen Kontrolle als technisch machbar erschienen – letztlich in einer kompletten künstlichen Welt, und gesteuert von – fast – allmächtigen künstlichen Intelligenzen wie etwa in der „Matrix“ Trilogie filmisch visualisiert.

Inzwischen hat es jedoch offensichtlich keine der bisherigen globalen Supermächte, oder auch keine der großen Regionalmächte geschafft, ein Orwell'sches Szenario komplett umzusetzen, auch wenn sich die großen Mächte aktuell gegenseitig verdächtigen, mit Hilfe von der durch sie dominant beherrschten KI/Künstlichen Intelligenz auch die Welt beherrschen zu wollen.

Natürlich stimmt dabei ein Einwand von Yuval Noah Harari: weder die spanische Inquisition noch das KGB haben - selbst wenn sie rund um die Uhr beobachtet haben - auch nur annähernd so viel Einblick in das Denken ihrer Opfer gehabt, wie demnächst unsere Biomarker über uns verraten werden - zumindest jene Daten, die wir über unseren Gehirnstrom freiwillig zur Verfügung stellen - und sei dies aus völlig harmlos naheliegenden Versicherungsgründen zu unserer Gesundheitsversorgung.

Vielleicht stimmt daher aber auch insgesamt eher eine andere Sichtweise auf Utopien und Dystopien: Henry Farrell sagt etwas sehr kluges in seinem Essay *"We live in Philip K. Dick's future, not George Orwell's or Aldous Huxley's"* in der „Boston Review“ (vom 16. Jänner 2018), in dem er ausführt, dass sich unzählige „pseudoreale“ und „reale“ Welten nebeneinander entwickeln und beeinflussen, zuletzt visualisiert in einem Remake von „Bladerunner“. Tatsache bleibt jedoch, dass auch in den Dystopien George Orwell's or Aldous Huxley's alle unsere Grund- und Freiheitsrechte in den klassischen Menschenrechten angesprochen sind, so wie wir sie in den letzten Jahrhunderten im Westen entwickelt haben und wie sie seit über einhundertfünfzig Jahren bei uns gelten.

Aber dennoch kam auch in den großen Geschichten beispielsweise George Orwell's oder Aldous Huxley's niemand auf die Idee, Rechte für künstliche Intelligenzen zu formulieren, oder aber jene Rechtsfragen zu präzisieren, wie lange denn unsere Gedanken in unserem Gehirnstrom unser Eigentum und frei sind, und ab wann unsere Gedanken gegen uns verwendet werden dürfen, falls wir im Netz eingehängt sind?

## Kapitel 2

---

### Zentrale Begriffe, Themenfelder und Fragen

Definitionen von KI/ Künstlicher Intelligenz sind entweder extrem komplex mit einer Unzahl an offenen Möglichkeiten und Bedingungen, oder aber pragmatisch kompakt. Als Beispiel hierfür verwenden wir daher in unserer laufenden Administration den sehr summarischen Ansatz der deutschen Wikipedia, wenn naturgemäß nur als Arbeitshypothese: *„Künstliche Intelligenz (KI, auch Artificielle Intelligenz (AI bzw. A. I.) ... ein Teilgebiet der Informatik, welches sich mit der Automatisierung intelligenten Verhaltens befasst. Der Begriff ist insofern nicht eindeutig abgrenzbar, als es bereits an einer genauen Definition von „Intelligenz“ mangelt. Dennoch wird er in Forschung und Entwicklung verwendet. ...im Allgemeinen bezeichnet künstliche Intelligenz den Versuch,*

eine *menschenähnliche Intelligenz nachzubilden.*“ ([https://de.wikipedia.org/wiki/Künstliche\\_Intelligenz](https://de.wikipedia.org/wiki/Künstliche_Intelligenz); letzter Zugriff am 1.1.2018).

In unserer Denkschrift verwenden wir den Begriff der KI vor allem für jene selbstmodifizierenden Programme, die auf der Basis neuronaler Netzwerke arbeiten (und die oft unter dem Begriff des „*Deep Learning*“ charakterisiert werden). *Dieser Fokus hat drei Gründe:* erstens erreichen selbstmodifizierende Programme, die auf der Basis neuronaler Netzwerke operieren, derzeit weltweit die erfolgreichsten Ergebnisse. Zweitens ist die Theorie neuronaler Netzwerke seit Jahrzehnten bekannt und die explosionsartigen Erfolge der KI der letzten drei Jahre bauen auf den hohen Rechenleistungen auf, mit der man inzwischen auch über tausend Lagen von neuronalen Netzwerken übereinander stapeln kann. Drittens ist es natürlich denkmöglich, dass andere Ansätze und Verfahren von KI umsetzbar sein können – was jedoch derzeit noch nicht konkret absehbar ist, wenngleich aktuelle Fortschritte in der Analyse von Neuronen durch genetische Verfahren große Fortschritte machen, was wiederum auf völlig neue neuronale Netzstrukturen hinweist, die in Zukunft modellierbar und programmierbar sein könnten.

Parallel publiziert das Wissenschaftsministerium den österreichischen „*Forschungsatlas*“ -für alle im Netz jederzeit einsehbar (<http://www.forschungsatlas.at/zukunftstechnologien/> Diese Forschungsarbeit wurde von *Policy Horizon Canada* in Auftrag gegeben und wurde in *MetaScan 3: Emerging Technologies April 2014 veröffentlicht*), und definiert dort unter „Schnittstellen“ wie folgt: *Biologisch erweiterte Sinne:* Basiert auf dem Prinzip der Gehirnevolution zur Erfassung einer einzigen Realitätskonstruktion, wobei nun mehrere lokale und entfernte Erfahrungen gleichzeitig übereinandergelegt stattfinden können, wodurch neue kognitive Erlebnisse möglich werden. Biologische Sinne können künstlich gesteigert und erzeugt werden und sich unterschiedlichen Stimuli anpassen sowie sich nach deren Maßgabe je nach Zweck ändern. „*Alle Sinne umfassende Multi-User-VR*“: <sup>[1]</sup><sub>[SEP]</sub>Eine alle Sinne umfassende Virtual Reality-Umgebung, mit der sich die Userin bzw. der User mittels direkter Gehirnstimulation verbindet. Durch die Anregung aller Sinne würden die Grenzen zwischen Realität und Fiktion verschwimmen. *Kontextbasierte IT:* Dies sind Computer, die sowohl ihre Umgebung spüren als auch auf sie reagieren können. Diese Geräte werden über Kontextinformationen zu ihrem Betriebsumfeld verfügen, die auf Regeln und Input von Sensoren basieren, und entsprechend reagieren können. Kontextbasierte Geräte könnten auch Annahmen über die aktuelle Situation von Nutzerin oder Nutzer treffen.

Der Forschungsatlas definiert weiters unter „*Neurotechnologie und Kognitive Technologie*“: *Gehirn-zu-Gehirn-Schnittstellen:* Es handelt sich um die hypothetische Umsetzung von Gehirn-Schnittstellen, die Gedanken, Empfindungen oder Impulse in digitale Signale übertragen und dabei die Daten im Gehirn der empfangenden Person rückumwandeln, um somit auf beiden Seiten eine Reaktion zu ermöglichen. Breit gefasst als Telepathie zu verstehen, könnten Gehirn-Schnittstellen Informationen von einer Person zur anderen transportieren, wobei das Internet lediglich eine vermittelnde Funktion erfüllt. Dadurch kann das empfangende Gehirn verhaltensbezogene Aufgaben ohne vorherige Schulung erfüllen. *Gehirn-Computer-Schnittstelle der nächsten Generation:* Hypothetische Schnittstellen zur Unterstützung, Erweiterung oder Reparatur menschlicher kognitiver oder sensomotorischer Funktionen sowie zur Kommunikation von Gedanken und Absichten an das Internet. *EEG Gehirn-zu-Computer-Schnittstellen:* Elektroenzephalografie ist und bleibt die einfachste Methode

zur Umsetzung von Gehirn-zu-Gehirn-Schnittstellen. Es ist das beste Tool mit räumlicher Auflösung zur Abbildung des aktiven Gehirns und ist darüber hinaus tragbar, nichtinvasiv und im Vergleich zu anderen Verfahren extrem kostengünstig.

Vor allem hier sind die Fortschritte der nichtinvasiven Intervention aus der US-Militärmedizin beeindruckend, wenngleich die EEG Technologie traditionell in Österreich entwickelt (Hubert Rohracher, et al.) und weltweit durch Giselher Guttmann erstmals als „Brainscanner“ erfolgreich eingesetzt worden ist - und nach wie vor beeindruckende Ergebnisse erzielt (ÖAW/Österreichische Akademie der Wissenschaften, Kommission für die Zusammenarbeit mit Dienststellen des BMLVS, Projektbericht 15, Psychologie und Landesverteidigung – Synergien aus anwendungsöffener Grundlagenforschung, Wien, Oktober 2013).

Kompilieren wir die Forschungsförder-Programme in Europa (EK und EDA), der USA (insbes. DARPA) und der NATO (jeweils soweit öffentlich) im Hochtechnologiebereich, so sind die Top- Prioritäten der Themenfelder evident:

- die Datenerfassung und deren Digitalisierung, mit der gesamten Sensorik, und den damit verbundenen Auswirkungen auf das gesellschaftliche Selbstverständnis,
- der gesamte Bereich der künstlichen Intelligenz/KI gemeinsam mit der Kooperation von Mensch und Maschine, dem „Human Machine Teaming“ und der Mensch/Maschine-Schnittstellen,
- sowie der Fächerübergreifende Bereich des „Cognitive Modeling of the Opponent“;

(Vgl.: <https://www.eda.europa.eu/what-we-do/our-current-priorities/research-technology>, oder die US –DARPA Programme unter <https://www.darpa.mil/our-research>).

Die Neurowissenschaften beschäftigen sich parallel dazu mit jenen Gehirnprozessen, die dem menschlichen Verhalten zu Grunde liegen, weshalb viele Autoren die Auffassung vertreten, dass sowohl Rechtswissenschaften wie auch Neurowissenschaften bestimmt sind als „natürliche Partner“, und berufen sich seit 2010 auf Goodenough und Tucker (<http://www.annualreviews.org/doi/abs/10.1146/annurev.lawsocsci.093008.131523>).

Die juristisch fehlende Grenze, in welchem Ausmaß Gehirnstromdaten auch im Falle einer aktiv gewollten Kommunikation durch Persönlichkeitsrechte geschützt sind, wird diskutiert unter „Kognitiven-“ und „Mentalen Rechten“ bzw. „Cognitive Liberty“, teilweise unter „Datensicherheit“ und „Privatheit/Privacy“. Sie verlangt sowohl eine menschenrechtliche Neufassung der Grund- und Freiheitsrechte als Schutznormen, als auch humanitär völkerrechtliche Normen insbesondere im „Cyberspace/Cyberwar“ einschließlich des Neutralitätsrechtes (Fichtenbauer et alii, bspw. in <http://scienceblog.sfu.ac.at/psychologie-heute-fragt-wie-lange-noch-sind-die-gedanken-frei-volksanwaltschaft-diskutiert-darueber-mit-justizminister-yuval-noah-hararis-homo-deu>).

Extrem sensibel ist der Einsatz von predictive AI/KI, nicht nur in Medizin oder im Internet, sondern in der Justiz bei Neurotechnologien, die ein zukünftiges Verhalten mit einer hohen Wahrscheinlichkeit voraussagen können. Verbindliche Definitionen fehlen, aber schon im Jahr 2013 war es möglich, dass man von 96 männlichen Strafgefangenen vor ihrer Entlassung eine klare Verbindung von neurophysischen Merkmalen und der Rückfallquote gemessen hat: Jene Strafgefangenen, die eine geringe Aktivität im Bereich des Anterior Cingulate Cortex (ACC) aufwiesen, hatten ein doppelt so hohes Rückfallrisiko innerhalb von vier Jahren wie ihre entlassenen Mitgefangenen mit durchschnittlichen Aktivitätswerten in der ACC Region (vgl. ua.: Aharoni ad Alii 2013, <sup>SEP</sup>Proceedings of the National Academy of Sciences of the United States of America, PNAS “neuro prediction of future re-arrest” <sup>SEP</sup>PNAS 2013; 110 (15): 6223-8).

Die juristische Einordnung von selbstständig handelnder KI/ Künstlicher Intelligenz wird weltweit heftigst diskutiert. Manche gestehen inzwischen der KI/ Künstlichen Intelligenz das Recht zu, etwa IPR/Intellectual Property Rights, zu erwerben, also „Geistiges Eigentum“ (wie bspw. Marken-, Muster oder Patentschutz, etc.), so etwa Ryan Abbott, immerhin Patentanwalt am United States Patent and Trademark Office (USPTO) sowie Professor für Recht und für medizinische Wissenschaften ([http://bclawreview.org/review/57\\_4/01\\_abbott/](http://bclawreview.org/review/57_4/01_abbott/)).

Wir sehen also sehr gut, dass der „Kampf gegen Killerroboter“, so wie er weltweit von NGO’s und Wissenschaftlern getragen wird in der Initiative „Campaign to Stop Killer Robots“ zwar nicht in die Leere, aber weit daneben geht, wenn er sich (nur) gegen autonome Plattformen richtet. Oder in anderen Worten: der wirklich gefährliche Killerroboter ist nicht die autonome Drohne – es ist der/die/das „Skynet“.

Machen wir aber dennoch der guten Ordnung halber einen Blick auch auf die Argumente und die Petition der Wissenschaftler der Initiative „Campaign to Stop Killer Robots“, zuletzt beim aktuellsten (asiatischen) AI/KI Kongress. Der Text lautet im Original:

“AN OPEN LETTER TO THE UNITED NATIONS CONVENTION<sup>SEP</sup> ON CERTAIN CONVENTIONAL WEAPONS<sup>SEP</sup>As companies building the technologies in Artificial Intelligence and Robotics that may be repurposed to develop autonomous weapons, we feel especially responsible in raising this alarm.

We warmly welcome the decision of the UN’s Conference of the Convention on Certain Conventional Weapons (CCW) to establish a Group of Governmental Experts (GGE) on Lethal Autonomous Weapon Systems.

Many of our researchers and engineers are eager to offer technical advice to your deliberations.

We commend the appointment of Ambassador Amandeep Singh Gill of India as chair of the GGE.

We entreat the High Contracting Parties participating in the GGE to work hard at finding means to prevent an arms race in these weapons, to protect civilians from their misuse, and to avoid the destabilizing effects of these technologies.

We regret that the GGE’s first meeting, which was due to start today (August 21, 2017), has been cancelled due to a small number of states failing to pay their financial contributions to the UN.

We urge the High Contracting Parties therefore to double their efforts at the first meeting of the GGE now planned for November. Lethal autonomous weapons threaten to become the third revolution in

warfare. Once developed, they will permit armed conflict to be fought at a scale greater than ever, and at timescales faster than humans can comprehend.

These can be weapons of terror, weapons that despots and terrorists use against innocent populations, and weapons hacked to behave in undesirable ways. We do not have long to act. Once this Pandora's box is opened, it will be hard to close. We therefore implore the High Contracting Parties to find a way to protect us all from these dangers."

Es ist daher auf jeden Fall gut – und wir können nun einerseits auch als Österreicher zufrieden sein, dass die Expertengespräche auf Regierungsebene nunmehr doch in Genf im November 2017 begonnen haben. Wir haben aber damit auch zumindest zehn grundlegende Fragenkomplexe zu beantworten, die wir zwar schon laufend aktuell diskutieren (*auch öffentlich, soweit sie die zivilen Fächer neben Ökonomie und Sicherheit vor allem in Medizin, Recht, Psychologie, Therapie, Sozial- und Kulturwissenschaften oder technische Grundlagenfächer erreichen - insbesondere mit Giselher Guttman im SFU – Scienceblog: [scienceblog.sfu.ac.at](http://scienceblog.sfu.ac.at)*), und die sich nicht länger verdecken lassen:

- mit entscheidenden Auswirkungen auf Österreichs Position als Neutraler in Europa und der Welt,
- unserer Sicherheit als Neutraler,
- und dem Standort von Wissenschaft, Forschung
- und unserer Wirtschaft :

1.) Wie kann eine UN Waffenkonvention, so wie sie seit dem 13. November 2017 in Genf beginnend auf Expertenebene diskutiert wird, überhaupt völkerrechtlich sinnvoll geregelt werden, wenn die globalen Spitzenexperten Recht haben: bekannt ist etwa Jürgen Schmidhuber, der basierend auf seiner Erfahrung als Entwickler der LSTM, der Long Short Term Memory Algorithmen - die heute alle Sprach- und Übersetzungsprogramme dominieren - uns daran erinnert, dass ein generelles Verbot, so wie es von der „Campaign to Stop Killer Robots“ gefordert wird, am Problem deshalb vorbei geht, weil eben KI/Künstliche Intelligenz schon überall existiert – und daher nicht nur beim Militär (oder Polizei oder den Geheimdiensten der Großmächte) und deren autonomen Plattformen, den Drohnen zu Lande, im und unter dem Wasser sowie über uns in der Luft und im Weltraum? Ein schlichter Blick in den oben dargelegten BMWFW Forschungsatlas sollte Diplomaten, Militär und Zivilisten gleichermaßen überzeugen.

2.) Wie kann eine UN Waffenkonvention, so wie sie seit dem 13. November 2017 in Genf beginnend auf Expertenebene diskutiert wird, überhaupt völkerrechtlich sinnvoll geregelt werden wenn die wichtigsten und gefährlichste autonomen Waffen nicht jene autonomen Plattformen und „Killerroboter“ sind, also die Drohnen zu Lande, im und unter dem Wasser sowie über uns in der Luft und im Weltraum, welche beispielsweise die „Campaign to Stop Killer Robots“ verbieten will, sondern viel aktueller und viel gefährlicher in der Form von jener waffenfähigen KI/Künstliche Intelligenz auftaucht und in beliebigen zivilen Anwendungen und Netzen, sowie auch als Folgegeneration von „KI der KI/Künstlichen Intelligenz“ als intelligente Programme im Netzkrieg, im Cyberwar, und in jenen Technologien, wie sie völlig offen beschrieben sind etwa im BMWFW Forschungsatlas? Diese „Cyber-Killer-Robots“ sind zwar Körperlos, aber zweifellos fähig, ganze Kontinente mit „Blackout-Schlägen“

lahm zu legen – etwas, was man auch den kampfstärksten „Drohnen Killerrobotern“ in Armeestärke kaum zutrauen wird.

3.) Wie kann eine UN Waffenkonvention, so wie sie seit dem 13. November 2017 in Genf beginnend auf Expertenebene diskutiert wird, überhaupt völkerrechtlich sinnvoll geregelt werden, wenn im Netz und daher in einem „Cyberwar“ die Rechte und Pflichten „nur“ der Neutralen auf Basis der Haager Landkriegsordnung geregelt sind ? Es existieren zwar seit Jahren breite wissenschaftliche Diskurse über Ethik im „Cyberwar“, und natürlich halten sich beispielsweise die Streitkräfte der NATO-Staaten an Prinzipien, so wie sie richtungsweisend im Tallinn Manual (*Tallinn Manual on the International Law Applicable to Cyber Warfare*) als Rechtsgutachten des NATO Cooperative Cyber Defence Centres of Excellence erstellt worden sind.

Wie jedoch derartige Rechtspositionen von den anderen Mitgliedern der UNO, vor allem von den rivalisierenden Supermächten im Sicherheitsrat gesehen werden, dies ist völlig unklar.

Österreich ist daher gut beraten, sehr schnell und im Zusammenwirken mit anderen Neutralen in und außerhalb der EU, insbesondere der Schweiz, sehr rasch Lösungen zu finden - für Neutrale im Cyberwar. Immerhin verlangt das geltende humanitäre Völkerrecht sehr klare und eindeutige Maßnahmen der Neutralen im Konfliktfall – und es ermächtigt die Konfliktparteien selbst tätig zu werden, wenn der Neutrale säumig ist wie unten in Punkt 5.) ausgeführt.

4.) Ebenfalls klar ist: wenn der Status der Neutralen im Cyberwar nicht zufriedenstellend abschließend geregelt werden kann, dann werden andere Mächte in diesen Neutralen-Status drängen. Weltkonzerne hätten schließlich auch gerne den völkerrechtlichen Status von Neutralen im Sinne der Haager Landkriegsordnung unter einer Genfer-Konvention, einer digitalen Genfer Konvention sozusagen, einem „Digital Geneva“: *Am 2. Februar dieses Jahres schlug Microsofts Präsident und Chief Legal Officer Brad Smith nicht weniger vor, als die völkerrechtliche Privilegierung der IT Konzerne als Neutrale in einem Cyberwar – ein bestechender Vorschlag, der vor allem das Selbstverständnis und die realen Machtverhältnisse auf den Konfliktfeldern des Netzes vor und in einem Cyberwar beleuchtet.*

Aktuell hat der Microsoft-Präsident seine Positionen international nochmals während der 5<sup>th</sup> GLOBAL CONFERENCE ON CYBERSPACE vom 20. bis 24. November 2017 in Delhi untermauert.

Fairerweise sollte man einräumen, dass zumindest „Neutrale Zonen“ im Netz, analog zu „Schutzonen“ im Humanitären Völkerrecht nicht denkunmöglich sind - etwa die Versorgung von Spitälern, um diese auch vor Bombardements und Blackout-Schlägen zu schützen?

5.) Die UN-Waffenkonvention, zu der die Beratungen in Genf begonnen haben, stellt als Rechtsfigur ein Rahmenabkommen bereit, welches die grundsätzlichen Zielsetzungen und Regeln eher abstrakt klärt, und sie bezieht sich dabei auf das IV. Abkommen zu den Genfer Konventionen, die den Schutz von Zivilpersonen in



Kriegszeiten rechtlich absichert. Als Neutraler haben wir nun mehrere sehr spezielle Probleme, sobald waffenfähige KI/Künstliche Intelligenz von einer der Konfliktparteien im Netz, d.h. im Cyberspace eingesetzt wird.

Erstens können wir natürlich nicht mit gutem Gewissen darauf warten, dass ein Krieg formell im Cyberspace mit einer Kriegserklärung beginnt. Die wenigsten Kriege auf unserer Welt wurden in den vergangenen Jahrzehnten formell erklärt. Dennoch gilt selbstverständlich für uns sowie für alle Konfliktparteien das Humanitäre Völkerrecht in Nachfolge des klassischen Kriegsvölkerrechtes, sobald ein bewaffneter Konflikt ausbricht, auch wenn es sich um keinen formell erklärten Krieg handelt. Dabei gibt es die bestehenden Regeln für alle Beteiligten, und natürlich auch für das Verhalten von Neutralen. Diese Regeln sind nicht immer ganz präzise formuliert, manchmal auch einander widersprechend, und natürlich auch gelegentlich unvollständig, da sich auch im bewaffneten Konflikt die gesellschaftliche und ökonomische Realität und die technologische Entwicklung oft genug schneller wandelt, als die darauf folgende Antwort mit einer völkerrechtlichen Normierung. Dies gilt insbesondere auch für den Netzkrieg, den Cyberwar, der momentan in Begriff ist, sich zu einem Ausmaß zu entwickeln, welches dem Kalten Krieg unseligen Gedenkens durchaus ähnlich wird. Allerdings gibt es auch dafür die schon erwähnten Rechte und Pflichten der Neutralen, die seit über hundert Jahren fein säuberlich und unbestritten festgelegt worden sind.<sup>[1][2][3][4][5][6][7][8][9][10][11][12][13][14][15][16][17][18][19][20][21][22][23][24][25][26][27][28][29][30][31][32][33][34][35][36][37][38][39][40][41][42][43][44][45][46][47][48][49][50][51][52][53][54][55][56][57][58][59][60]</sup>

Die Haager Landkriegsordnung (Anlage zum IV. Haager Abkommen), als wichtigster Teil der Haager Abkommen, ist seit über hundert Jahren gültiges Völkerrecht, genau genommen Völker-Gewohnheitsrecht.<sup>[1][2][3][4][5][6][7][8][9][10][11][12][13][14][15][16][17][18][19][20][21][22][23][24][25][26][27][28][29][30][31][32][33][34][35][36][37][38][39][40][41][42][43][44][45][46][47][48][49][50][51][52][53][54][55][56][57][58][59][60]</sup>

*Die Anlage zur Haager Konvention „betreffend die Gesetze und Gebräuche des Landkriegs“ enthält (in den Fassungen von 1899 und 1907 in 60 beziehungsweise 56 Artikeln) die eigentlichen Festlegungen zu den Regeln und Gebräuchen des Landkrieges, und regelt als „Haager Abkommen betreffend die Rechte und Pflichten der neutralen Mächte und Personen im Falle eines Landkriegs“. Dies klingt formal, und es ist auch rechtlich formal, allerdings durchaus anwendbar auf den Netzkrieg, den Cyberwar:*

*Das „Haager Abkommen betreffend die Rechte und Pflichten der neutralen Mächte und Personen im Falle eines Landkriegs“ stellt nämlich ganz klar fest, dass der Neutrale zu unterbinden hat, dass eine der beteiligten Konfliktparteien eine eigene technische Kommunikationsanlage auf dem Boden des Neutralen für sich im Konflikt nutzt. Damit ist auch eine weitere Konsequenz glasklar: Österreich hat zu unterbinden, dass eine der Konfliktparteien ihre eigenen technischen Kommunikationsanlagen auf dem Boden Österreichs als eines Neutralen für sich im Konflikt nutzt.*

Wenn also ein Server einer der Konfliktparteien, nehmen wir einmal rein fiktiv an, für ein chinesisches unterstütztes Nordkorea, im Großraum von Wien, nahe Graz oder bei Linz einen illegalen Server verdeckt für Netzkriegsaktionen nützt, so hat Österreich dies zu unterbinden. Selbstverständlich gilt dies auch für den umgekehrten hypothetischen Fall, wenn ein koreanisch-japanisch-amerikanisch gestützter Server illegal gegen Nordkorea-China kriegerische Handlungen setzt.



6.) Nun sind zwei weitere Rahmenbedingungen ebenfalls klar, nämlich die technischen und die juristischen Möglichkeiten: Illegale Server können sich natürlich – elektronisch – im Netz tarnen und verstecken, zumindest für eine gewisse Dauer der Aktivitäten. Aber selbstverständlich können sie entdeckt und aufgeklärt werden. Das beherrschen ebenfalls alle ernst zu nehmenden Industriemächte, also in unserem hypothetischen Fall die Amerikaner mit den Südkoreanern und den Japanern, und selbstverständlich auch die Nordkoreaner und die Chinesen. Es beherrschen auch wir in Österreich schon recht passabel. Wenn Sie derzeit etwa schattig-spacig-grüne Plakate des Bundesheeres am Straßenrand sehen, so werden Sie lesen, dass unser Heer damit wirbt „*Die Bedrohungen ändern sich – wir uns auch*“ und damit die Kräfte für eine Netzkriegsführung durchaus zu recht präsentiert. Das Regierungsprogramm legt daher konsequenterweise die Schwergewichte auf Cybersicherheit, und die „Leuchtturmprojekte“ des BMLV sehen ein Cyber-Zentrum gemeinsam mit BMI und BKA vor. Dies entspricht der zweiten Rahmenbedingung unserer österreichischer Rechtsordnung, zumindest im Frieden: Wenn ein illegal agierender Server feindselige Akte der Netzkriegsführung setzt, so werden im Friedensfall entweder die betroffenen ausländischen Mächte, wahrscheinlich im Wege der diplomatischen Vertretungen unsere Justiz informieren, die dann entsprechende Maßnahmen durch die Polizei veranlasst.

Dasselbe gilt, wenn die Cyber-Kräfte des Bundesheeres eine illegale Aktivität entdecken: Dann werden diese Erkenntnisse des Heeres an Justiz und an Polizei übergeben zum Vollzug im Rahmen der bestehenden Rechtsordnung. Unangenehm wird nur ein anderes Szenario, welches sich im Schatten der Kriegsdrohungen rund um die koreanische Halbinsel unangenehm aufbaut: Was passiert, wenn eine der Konfliktparteien nicht warten will, oder nicht warten kann, bis der illegale Server am Rechtsweg unter Friedensbedingungen exakt identifiziert und dann auch abgestellt wird? Was passiert, wenn in einem rasant eskalierten kriegerischen Konflikt eine der beiden Konfliktparteien sich auf das „*Haager Abkommen betreffend die Rechte und Pflichten der neutralen Mächte und Personen im Falle eines Landkriegs*“ beruft, und selbst tätig wird, wenn wir in Österreich nicht schnell genug tätig werden können – oder besser gesagt nicht tätig werden „dürfen“, weil das humanitäre Völkerrecht für den Neutralen im Netzkrieg/Cyberwar noch nicht aktuell interpretiert werden konnte, während gleichzeitig schon die großen IT Weltkonzerne für sich eine digitale Genfer Konvention urgieren?

7.) Ganz besonders unangenehm wird es, wenn wir mit einer „KI/Künstlichen Intelligenz der Künstlichen Intelligenz“ bei Netzattacken rechnen müssen. Der nächste Netzkrieg, ein Cyberwar, wird natürlich mit künstlich intelligenten Programmen geführt, also einer Art von „Cyber-Killer-Robots. Man kann sich unschwer ausmalen, was passiert, wenn dabei Trojaner und Viren eingesetzt werden, die ihrerseits künstlich intelligente Programme schaffen, also „KI/Künstlichen Intelligenz erzeugt durch KI/Künstlichen Intelligenz“ und so den Ausgangspunkt und die Urheber fast perfekt tarnen und vielleicht werden sie ja richtig erfinderisch.

Aber auch wenn uns ein Netzkrieg, ein Cyberwar mit „Cyber-Killer-Robots“ erspart bleiben soll: KI der KI kommt natürlich auf jeden Fall, was uns natürlich gleich zu dem nächsten Rechtsproblem führt, und zwar auch dann, wenn die Entwicklungen der KI's auf den Schultern der KI's gar nicht feindselig sondern nützlich sind: denn wem gehören eigentlich die Erfindungen, die unsere KI's machen? Und wer ist für den Schaden verantwortlich?

Wir sind erst am Beginn der Diskussion - weniger in den ethischen Diskursen weltweit, sondern zur notwendigen rechtlichen Umsetzung. Dies trifft am dringendsten das Militär wegen des humanitären Völkerrechtes und dem weltweiten Wettlauf in den disruptiven Technologien, und in der Medizin, weil die meisten persönlichkeitsverändernden Technologien zu aller erst als Heilung entwickelt und eingesetzt werden wird – beginnend von Wiederherstellungsverfahren nach Gehirnverletzungen.

Zur rechtlichen Umsetzung des Status von „KI/Künstlicher Intelligenz geschaffen durch KI/Künstlichen Intelligenz“ darf nochmals auf Ryan Abbott’s Position verwiesen werden. Bemerkenswert an diesem Vorstoß ist ja auch die enge Verbundenheit von Ryan Abbott mit medizinischen Patenten, die – wenn die Rechtsordnung seinen Anregungen folgen sollte – in Zukunft auch von KI-Robotern gehalten würden, falls diese die Erfinder waren. Wer dies nochmals genau nachlesen möchte: [http://bclawreview.org/review/57\\_4/01\\_abbott/](http://bclawreview.org/review/57_4/01_abbott/)

8.) Damit sind wir wieder bei unseren aktuell (in der ÖAW sowie der Wissenschaftskommission und in unserem [scienceblog.sfu.ac.at](http://scienceblog.sfu.ac.at)) diskutierten zentralen Fragen: Wie behandeln wir einerseits jene autonomen KI-Systeme, die ihrerseits selbst von autonomen KI-Systemen geschaffen worden sind, wie etwa im Cyber-War der rivalisierenden Großmächte und ihrer Stellvertreter von Korea bis in den Nahen Osten, und im Wettlauf um den Cyber-Space? Diese „autonomen KI-Systeme, die ihrerseits selbst von autonomen KI-Systemen geschaffen worden sind“ stellen das eine Grundproblem dar, welches wir sehr rasch lösen müssen.

9.) Andererseits geht es aber schon heute sehr konkret auch um unsere persönlichsten Daten (nicht nur, aber speziell auch) wenn wir „aus Versehen“ abgegriffen werden, nicht nur von erfinderischen KI-Robotern, sondern sehr unmittelbar beispielsweise aus unseren Smart-Devices und den Datenbrillen, die wir mit Gehirnstrom steuern werden, so wie dies von Facebook, Apple oder Samsung auf der Basis von EEG-Technologie angekündigt worden ist. Was passiert dann rechtlich und faktisch, wenn unsere „zufällig“ erfassten Gehirnstrom-Daten nicht nur bei den großen Firmen oder bei Verbrechern landen, sondern natürlich auch bei kontrollwütigen Bürokraten, bei Zensoren oder bei einer Gedankenpolizei nach Orwell’schem Vorbild?

Was passiert eigentlich verfahrensrechtlich, wenn dann unsere „zufällig“ erfassten Gehirnstrom-Daten vor jedem Gericht gegen uns selbst verwendet werden könn(t)en, weil de facto „unser Gehirn uns selbst verrät“ – und dabei sprechen wir noch gar nicht von Gehirn-Implantaten wie dem „Neuronal Lace“ und seinen Verwandten. Hier stellen sich auch alle grundlegenden Fragen zu den Grund und Freiheitsrechten, wie wir sie im Diskurs von Volksanwalt Peter Fichtenbauer mit Justizminister Wolfgang Brandstätter behandelt sehen: *„was (aber) machen wir morgen, wenn es gelingt, dass die Gehirnstromdaten die Grenzen unserer Persönlichkeit angreifen und auflösen? Die Bindung des Gendarmen, des Richters im Jahr 1867 an die Prinzipien des Staatsgrundgesetzes, also die Freihaltung von unakzeptablen Eingriffen, sind heute der biochemischen und technischen Bemächtigungsmöglichkeit über den Menschen gewichen“.*

10.) Und nicht zuletzt die Frage, wie gehen wir in Österreich insgesamt mit den Grenzfragen um, die sich ja nicht nur in den DARPA Forschungsthemen spiegeln, sondern die evtl. ab dem kommenden Jahr den größten Einzelbereich der europäischen Forschung ausmachen, und in der europäischen Verteidigungsforschung den EDA

Prioritäten folgend allen Mischformen über die Fakultätsgrenzen hinweg folgend beginnend vom Human-Enhancing bis zum Human-Machine-Teaming?

## Kapitel 3

### Wie „intelligent“, wie „einsichtig“ oder „weise“ ist Homo Sapiens? Und wie „intelligent“ ist KI/Künstliche Intelligenz?

Darauf gibt es so wahrscheinlich so viele Antworten wie Menschen.

Verhaltenspsychologisch sind etwa die kognitiven Fähigkeiten unserer evolutionär engsten Verwandten, der Primaten, mit denen von kleinen Kindern gut vergleichbar (rezent etwa Felicitas Auersperg in <https://www.sfu.ac.at/de/news/psy-felicitas-auersperg-das-merkwuerdige-verhalten-von-schimpanzen-in-kinderkleidung/>).

Aber es geht natürlich nicht nur um kognitive Fähigkeiten im engeren Sinne - um beispielsweise mit Immanuel Kant zu sprechen: *Das Erstaunen über den gestirnten Himmel über mir...und über das Sittengesetz in mir ...* mit den ewigen Fragen: *Woher komme ich? ... wohin gehe ich? ... und worauf dürfen wir hoffen?* Diese Fragen stellt sich wohl jeder Mensch – und vielleicht auch noch zu unseren Lebzeiten die ersten künstlichen Intelligenzen.

Zählen wir ergänzend zu den a-priori vorstellbaren Größen, also Raum, Zeit, Kausalität und Modalität auch jene Phänomene, deren Existenz sich uns nur indirekt erschließt – durch Erfahrung etwa, oder durch die Phänomenologie in ihren Auswirkungen, so sind unbestreitbar etwa die Kategorien Platons für jeden Menschen sehr unmittelbar erfassbar: das Wahre, Gute und das Schöne - obwohl wir nur sehr begrenzt in der Lage sind verbindlich zu definieren, was denn nun bspw. alles „schön“ ist.

Die Beobachtung von möglichen, langweilig zu beobachtenden Rückkopplungsprozessen führt uns in jene Diskussionskapitel, in dem wir *kulturell erworbenes implizites Wissen als Gesellschaft in unseren sozialen Interaktionen verwenden: für Ergebnisse, die sich jeder individuellen Planbarkeit entziehen*. In diesem Sinn ist daher jeder funktionierende Markt, sowohl im Austausch von Ideen, als auch im Austausch von Leistungen, ein Erkenntnis-gewinnender Prozess, der unendlich mehr Information enthält, als es jemals ein einzelnes Gehirn erfassen, planen oder steuern könnte. Das ist nicht neu, und es macht alle unsere menschlichen Kulturen aus, die uns umschließen, und die niemals von Einzelmenschen in ihrer Tiefe und ihrer Breite verarbeitet oder gar geplant werden könnten – nicht einmal in den totalitärsten aller Diktaturen.

Wo aber stehen heute selbstlernende KI/Künstliche Intelligenzen? Künstliche Intelligenzen kennt jeder, zumindest in den dramatischen Bildern zahlloser Romane und der Science Fiction Filme. Sie haben unsere Vorstellungen geprägt - von der „Space Odyssey“ über das „Sky Net“ im „Terminator“ bis zu Internetwelten der „Matrix“. Auch Mensch/Maschine-Schnittstellen, einschließlich der BCI/Brain Computer Interfaces finden wir zwischen Cyborgs, „Iron Man“ und „Avengers“ oder „Avatar“, in denen sich unserer Geist und unsere Individualität mischt mit anderen und mit ganzen, Planeten umspannenden Netzen.

Die gesellschaftlichen und auch die wissenschaftlichen Begleitdebatten schwanken dabei stets zwischen Beschwichtigung und Alarmismus - also zwischen der beruhigenden Behauptung, dass es sich lediglich um Phantasie-volle „Science Fiction“ handelt, und den Weckrufen, dass wir es schon längst mit „Science“ zu tun haben, und nicht mehr mit bloßer „Fiction“.

„Dumm wie ein Sieb“ sei die KI/Künstliche Intelligenz, so hört und liest man allenthalben (beispielsweise im deutschen Nachrichtenmagazin „Der Spiegel“, gleich in der zweiten Ausgabe zu Beginn des Jahres). Viele kluge Gelehrte weisen darauf hin, das selbst dreijährige Kinder besser und schneller geeignet sind, Ähnlichkeiten zu erkennen – auch wenn unsere Dreijährigen normalerweise nicht in der Lage sind, strategische Bomber auf schwankenden Flugzeugträgern zu landen, oder Hautkrebs und Herzrhythmusstörungen an den medizinischen Geräten besser zu erkennen, als auch ziemlich erfahrene Fachmediziner. Gleichzeitig ist das weltweit am meisten verbreitete Werk ausgerechnet die Betrachtung eines Historikers, Yuval Noah Harari, Militärhistoriker aus Jerusalem, der in seiner Universalgeschichte „Homo Deus“ den Bogen spannt vom Beginn der Menschheit bis zu den Möglichkeiten eines „Transhumanismus“, von dem wir nicht wissen, wohin er uns führen wird, sobald die Rechner wesentlich mehr wissen und können als wir Menschen, und vielleicht selbst so etwas wie ein reflexives Bewusstsein entwickeln. Ob sie diese Selbsterkenntnis schaffen, ist heftig umstritten.

Völlig klar absehbar ist jedoch, dass wir menschliches Wissen und Bewusstsein auf Gehirnstrombasis vernetzen und demnächst transferieren können, im Austausch mit anderen Menschen und ins Netz auch über Weisheit, Schönheit und Stärke, oder zumindest dem, was wir als Weisheit, Schönheit oder Stärke empfinden und zu erkennen glauben ... und damit vielleicht doch auch „Die letzten Tage des Todes“ begonnen haben werden, wie dies Yuval Harari in seinem „Homo Deus“ diskutiert.

Damit sind wir auch schon an einer anderen, sehr stimmungsvollen und eindringlichen Betrachtung des Endes unseres Lebens. Heuer ist es gerade mal 60 Jahre her, dass Ingmar Bergman weltberühmt wurde mit seinem Film „Das siebente Siegel“, in dramatischen Symbolen und Bildern, und parallel zu den davor noch nie gesehenen Bildern eines erstmals von Menschenhand in den Weltraum beförderten Wesen, der Hündin Laika, unmittelbar nach dem Sputnik als erstem Satelliten. In Schwarzweiß gehaltenen Bildern und im Stile eines Mysterienspiels nach dem Vorbild eines mittelalterlichen Totentanzes zweifelt im „ siebenten Siegel“ ein heimkehrender Kreuzritter an seinem Glauben, gezeichnet von zahllosen Kämpfen, der versucht ritterlich Frauen zu schützen vor Verbrechern, und der auf den personifizierten Tod trifft. Dieser will ihn mitnehmen, sofort. Er ist jedoch einverstanden die Lebensspanne des heimkehrenden Kriegers und Ritters solange aus zu dehnen, als dieser ihn im Schachspiel besiegen kann.

Das heurige 60-Jahre-Jubiläum des „Siebenten Siegels“ ist nicht nur eine Gelegenheit, die eindrucksvollen Bilder Ingmar Bergmans zu dessen hundertsten Geburtstag wieder auf uns einwirken zu lassen. Es ist auch der Anlass für eine Positionsbestimmung, wo denn der personifizierte Gevatter Tod im Schachspiel steht, und wo wir, als fühlende und intelligente Wesen denn im Vergleich zu den aufkommenden künstlichen Intelligenzen stehen? Unbestritten greift Ingmar Bergmann auf Motive zurück, die sehr tief in unserer Kultur verankert sind. Der Tod als „Gevatter Tod“ in unseren Märchen , der sich solange zurück hält, wie noch eine Kerze brennt, und nur

sichtbar für die Hauptfigur aus Grimm`s Märchen „Der Gevatter Tod“ ist nicht per se böswillig – und wird von einem „armen, verzweifelten Mann“ gesucht für sein dreizehntes Kind als Taufpaten, als einen „Gevatter“, nachdem der Mann „den lieben Gott“ als Taufpaten, als „Gevatter“ abweist („du gibst den Reichen und lässt den Armen hungern“), sowie den Teufel („du betrügst und verführst die Menschen“) und erst denjenigen „der alle gleich macht“ akzeptiert.

Natürlich sind viele Darstellungen auch dunkler. Wir kennen den personifizierten Tod als „Schnitter“, als „Grim Reaper“, oder im „Jedermann“, usw., sowie im spätmittelalterlich populären Totentanz, und dabei selbstverständlich auch den Tod personifiziert als Schachspieler (auch in den Landen Ingmar Bergmanns, etwa in der Kirche von Täby, gemalt von Albertus Pictor im 15. Jahrhundert).

Jetzt, im 21. Jahrhundert, sieht jedoch Yuval Noah Harari „*Die letzten Tage des Todes*“ herauf ziehen in seinem Weltbestseller „*Homo Deus*“. Nein, nicht deshalb, weil wir Menschen plötzlich unsterblich geworden sind. Allerdings, so Yuval Noah Harari, „huldigten etwa Christentum, Islam oder Hinduismus stets etwas, das oberhalb oder jenseits des irdischen Daseins stand ... (und) einige waren geradezu verliebt in Gevatter Tod“ - und Yuval Noah Harari fährt fort: „In Wirklichkeit aber sterben Menschen nicht, weil eine dunkel gewandete Gestalt sie an der Schulter packt“, sondern weil vitale Systeme unsere Körpers versagen, manchmal natürlich auch wegen gewaltsamer Einwirkung von außen.

Lassen wir an dieser Stelle vorerst einmal beiseite, wie wir mit unserem heutigen Wissen vielleicht auch die letzten Fragen des Menschen neu aufrollen könnten: *Woher komme ich? ...wohin gehe ich? ...und worauf dürfen wir hoffen?*...um beispielsweise mit Immanuel Kant zu sprechen. Unbestritten bleibt jedoch, dass sich vor gerade einmal 60 Jahren ein personifizierter Tod mit einem Menschen gemessen hat, mit einem heimkehrenden Krieger, einem Ritter, im Wettkampf in einem alten Strategiespiel das seit Jahrhunderten eingebettet in unsere Kultur ist, dem Schach– und wir alle diese filmische Visualisierung als erschütterndes existentialistisches Kunstwerk, und als intelligent-verständnisvoll und tiefsinnig erleben konnten.

Wenn wir uns nun als Menschen selbstkritisch und ehrlich (ver)messen wollen, dann ist nicht nur die Frage offen, ob – dereinst vielleicht in einem „*Transhumanismus*“ - unsere physisch irdische Identität in ein Netz geladen werden könnte, und wir dadurch faktisch unsterblich würden. Wir sollten uns gerade jetzt auch ehrlich fragen, woran wir denn bisher beispielsweise so etwas wie „Einsicht, Erkenntnis und Intelligenz“ gemessen haben (wobei wir uns vorweg noch kurz von den Messungen der Verteilungskurven kognitiver Fähigkeiten in standardisierten Intelligenztests lösen wollen).

Selbstverständlich hätten vor 60 Jahren wohl die allermeisten Menschen auf beiden Seiten des Eisernen Vorhanges angenommen, dass eine unglaublich gebildete, einsichtsvolle Intelligenz gegen sie wirkt, wenn diese ihre amtierenden Großmeister im Schach schlagen kann. Stellen wir uns einen Wettkampf vor 60 Jahren vor, eine internationale Schachmeisterschaft, in der das Prestige der Supermächte gefordert ist, im ersten Jahr des Weltraumfluges mit dem Sputnik (1957), und erlauben wir einem Team von Großmeistern alle nur denkbaren Hilfen, und mit Beraterinnen und Beratern ihrer Wahl auf einer Seite des Vorhanges zu sitzen, und auf der

anderen Seite des Vorhanges eben der oder die Gegenspieler, offensichtlich unschlagbar intelligent und strategisch einsichtig.

Noch zu Zeiten des kalten Krieges hätte das wohl jede der Supermächte als ein unerreichbares Beispiel an kulturell konnotierter Intelligenz gewertet.

Wie schon angerissen, messen wir weltweit kognitive Fähigkeiten in Häufigkeitsverteilungen der Intelligenztests (mit allen Einschränkungen, aber auch mit getesteter Validität der Aussagekraft dieser differentialpsychologischen Diagnostik), und es ist für uns darüber hinaus auch implizites Wissen als „Erkenntnis“, und manchmal auch als Weisheit, als Schönheit oder als innere Stärke zumindest erfahrbar, wenn auch nicht immer exakt messbar – im Erwachen unseres Geistes bis zu unserem Abschied aus unserem Leben, und jeder wird wohl die Erbauer von Stonehenge, die Architekten der Pyramiden oder gotische Dom-Baumeister als „intelligent“ bezeichnen, auch, weil sie hochwertigstes Fachwissen und implizite Informationen, also ein „Wissen“ etwa über ästhetischer Kategorien, wie „Schönheit“ vereinen konnten.

Wir erkennen implizites „Wissen“ sogar empirisch messbar bei quantitativen Systemübergängen, wiewohl oft erst im Nachhinein: wie viele Sandkörner bilden einen Haufen – mit und ohne Sieb, und wie viele Haare formen einen Schwanz? Implizite Informationen sind auch in einem anderen universellen Phänomen sehr gut gespiegelt – etwa in epidemiologischem Wachstum – und dort recht gut berechenbar in Diffusionsgleichungen: Wir kennen dies aus Wachstum in ökologischen Nischen und in Marktnischen. Es ist meist gut berechenbar in Diffusionsgleichungen bei Bakterienkulturen auf Agar-Agar, bis hin zu der ökologischen Abhängigkeit der Fressfeinde von ihren Beutetieren, oder in der explosionsartigen Ausbreitung des Lebens auf unserem Planeten.

Hierzu gehören auch alle Modelle von langwelligen Entwicklungen: Etwa die großen Substitutionswellen der Energieversorgung, oder die uns heute bestimmenden Schlüsseltechnologien – also etwa die Substitution von Pferdekarren durch Autos, der Ersatz von Kohle durch Erdöl, und inzwischen von Erdöl durch die Generatoren moderner Windkraftwerke und durch Solarzellen mit steigendem Wirkungsgrad, usw.

In diesem Sinn ist daher jeder funktionierende Markt, sowohl im Austausch von Ideen, als auch im Austausch von Leistungen, ein Erkenntnis-gewinnender Prozess, der unendlich mehr Information enthält, als es jemals ein einzelnes Gehirn erfassen, planen oder steuern könnte. Das ist nicht neu, und es macht alle unsere menschlichen Kulturen aus, die uns umschließen, und die niemals von Einzelmenschen in ihrer Tiefe und ihrer Breite verarbeitet oder gar geplant werden könnten – nicht einmal in den totalitärsten aller Diktaturen.

Wo aber stehen heute selbstlernende KI/Künstliche Intelligenzen? Sehen wir uns als Beispiel die Entwicklung zweier Kulturen an, die wohl alle kennen, und die seit Jahrhunderten eingebettet sind in unsere Vorstellung von „Intelligenz“:

zwei Systeme mit klaren Prinzipien, und vor allem mit überschaubar widerspruchsfreien Regelsätzen. Wir waren wir uns dennoch seit Jahrhunderten absolut sicher, dass die Meisterung dieser Systeme viel mehr verlangt, als die bloße Kenntnis der Regeln und den Austauschprozess mit dem Gegenüber.

Betrachten wir daher zuerst das „Spiel der Könige“, das Schach.

Bei der Klassifikation der Spielstärken halten wir uns traditionell an einen weltberühmten amerikanischen Ungarn, Arpad Elöl, mit den nach ihm benannten ELO Punkten für das Schachspiel: *Es handelt sich dabei um Ranking-Verfahren*, das auf Schach genauso wie auf eine Fußball-WM, auf Tischtennis oder auf Klassen von Medizinpreisträgern, Juristen, Sängern, Hauben-Restaurants, auf Wissenschaften oder Generäle angewandt werden kann – und auch auf Computer untereinander, und wir können es verbinden mit den Analysen eines anderen weltbekannten Ungarn, mit Laszlo Merö (*„Habits of mind, The power and limits of rational Thought“*, 2002).

Folgen wir dieser Idee der Ranking – Verfahren, und insbesondere Laszlo Merö in seinem Ansatz, dann können wir in den Strategiespielen etwas finden, was er die „Tiefe“ eines strategischen Spieles genannt hat. *Wettbewerbskategorien werden dabei dadurch definiert, dass Klassen von Teilnehmern nach ihrer Fähigkeit zusammengefasst sind, wenn diese in der Lage sind, zumindest 75% aller anderen Teilnehmer zu schlagen.*

Mit dieser Kategorisierung ist es denklogisch, dass wir Teilnehmerklassen für jede Art von Wettbewerben formulieren können, egal ob dieser Bewerbe auf sehr präzisen Messergebnissen beruht wie bei Autorennen, oder auf sehr subjektiven Bewertungen, wie etwa beim Vorspielen mit Instrumenten vor großen Musikorchestern. Betrachten wir Wettkampf-Klassen bspw. bei verhältnismäßig unbestritten messbaren Leistungen, bei Laufzeiten von Sportlern, bei Fahrzeiten in Regatten oder auch in der Kapazität der Informationsverarbeitung etwa in den (speziellen Teilgebieten von kognitiven Fähigkeiten bei) Intelligenztests: Wir wissen auch, dass etwa die Differenz zwischen einem IQ von 94 auf einen IQ von 130 annähernd den Unterschied ausmacht von ca. dem doppelten Leistungsvermögen in einer äquivalenten Informationsverarbeitung. Wir alle wissen auch sehr praktisch, dass wir bei Übermüdung oder unter Alkohol und Drogen-Einfluss diese speziellen kognitiven Fähigkeiten unserer (hoffentlich gut entwickelten) „Intelligenz“ auch in diesem Ausmaß rasch beeinträchtigen können: in etwa von „teilweise hochbegabt“ herunter auf „teildebil“, und unfähig, auch nur das eigene Auto zu öffnen.

Es ist also nachvollziehbar, wie wir etwa auch im Schach, vom Laien bis zum Großmeister, klassifizieren können: indem man sie gegeneinander antreten lässt. Es werden sich aus der Evolution der Wettkampf-Ergebnisse heraus Klassen von Meistern ergeben – und man wird eben Großmeister dadurch, dass man andere, ebenfalls meisterlich spielende Mitbewerberinnen oder Bewerber zumindest meistens besiegen wird – und vor diesem Verständnis macht es Sinn, wenn eine personalisierte Wesenheit wie der Gevatter Tod aus den Märchen der Gebrüder Grimm und den spätmittelalterlichen Wandmalereien sich doch gelegentlich auf eine Partie Schach mit einem Menschen einlässt - so auch der Tod aus dem siebenten Siegel von Ingmar Bergmann vor 60 Jahren.



Selbstverständlich ist es nicht ganz eindeutig, was wir in den Wettbewerbsklassen messen: wir erfassen damit ja nicht nur kognitive Dimensionen des impliziten Wissens, sondern auch Faktoren von physischer und psychischer Widerstandsfähigkeit sowie zahllose Bedingungen eines umgebenden Systems. So sagt etwa eine Entscheidung nach einem Wettbewerbsverfahren über einen „Oscar“ Filmpreis oft mehr aus über die Jurorinnen und Juroren der AMPASS, der „Academy of Motion Pictures Arts and Sciences“ sowie über die Werbebudgets der Filmindustrien zwischen Bollywood und Hollywood – aber immerhin, es gibt damit ein weltweit als maßgeblich beachtetes Wettbewerbsverfahren.

Beobachten wir daher die Kultur der Strategiespiele auch im Kontext der großen strategischen Denkschulen, die das moderne strategische Denken beherrschen. Sie alle lassen sich – zumindest derzeit noch – nicht in programmierbare Algorithmen fassen, wenn es um strategische Entscheidungen in einem komplexen Umfeld geht. Wir blicken dabei auf Clausewitz als dem weltweit am meisten beachteten neuzeitlichen strategischen Denker genauso wie auf zahllose Inputs aus allen Kulturen, von Niccolò Machiavelli aus der europäischen Renaissance, über Persien bis Indien, und weiter auf Sun Zi (Sun Tsu), als dem sicher am meisten zitierten strategischen Denker des chinesischen Kulturkreises. Er hat als Zeitgenosse fast parallel zu Thucydides gewirkt, der bekannt ist mit seinen Werken über den peloponnesischen Krieg, also der Epoche zwischen den klassischen griechischen Staaten und dem Großreich der Perser und vom persischen Schah stammt ja auch der Begriff „Schach“, wodurch es eben zum „königlichen Spiel“ wurde.

Daneben steht natürlich auch die psychologische Faszination des Schach. Manche werden noch die „*kleine Schachnovelle*“ von Stefan Zweig kennen, sein letztes Werk aus dem brasilianischen Exil, wo eine Persönlichkeitsspaltung eines legendären Großmeisters die Hauptrolle spielt.

Es war auch niemand geringerer als der Naturforscher, Philosoph und einer der Gründungsväter der USA, Benjamin Franklin, der in einem 1779 geschriebenen Werk über „*The Morals of Chess*“ nachgedacht hat. Es handelt sich dabei um einen philosophischen Aufsatz, in dem Benjamin Franklin einen Vergleich zwischen den Spielregeln des Schachs und dem sittlichen Handeln des Menschen aufstellt. Selbstverständlich sind moderne Schachcomputer nicht speziell darauf optimiert, sich an diese sittlichen Regeln von Benjamin Franklin halten. Sie legen daher leider auch die von uns allen, und von Benjamin Franklin so geschätzte „*hochherzigen Höflichkeit*“ nur dann an den Tag, wenn wir sie darauf nachprogrammieren.

Was ein Schachcomputer allerdings besser kann, ist auch klar: blindes Rechnen, dies aber in höchster Geschwindigkeit. Echte Herausforderer für menschliche Großmeister wurden sie natürlich nicht vor 60 Jahren im Jahr des Sputnik und des „Siebenten Siegels“, sondern erst in den Jahren rund um den Fall des Eisernen Vorhangs und der Berliner Mauer, 1989, als IBM mit „*Deep Thought*“ gegen Kasparow angetreten ist. Als dann ein Jahrzehnt später „*Deep Blue*“ endgültig die Oberhand gewinnen konnte, 1997, war „*Deep Blue*“ in der Lage, ca. 200 Mio. Positionen pro Sekunde zu evaluieren.

Heute wäre dies nur mehr eine mäßige Leistung, allerdings – und dies ist von elementarer Wichtigkeit – wir konnten noch nachrechnen, wie der Computer zu seinem Ergebnis gekommen ist.



So wie das Schach von Persien aus die westliche Kultur geprägt hat, so hat auf der anderen Seite das Go, ursprünglich aus China kommend, in der gesamten konfuzianischen Welt seinen Siegeszug geschafft und ist daher nicht nur in der chinesischen Kultur, sondern auch in Japan und in Korea höchst populär. Werfen wir daher auch einen Blick auf die Ränge von Go-Spielern nach denselben oben angeführten Prinzipien in „Wettbewerbsverfahren“, so sehen wir, dass es mehr Wettbewerbsklassen als im Schach gibt, was dadurch entsteht, dass unser Problem des impliziten Wissens schon viel früher auftritt.

Schach umfasst zwar eine den menschlichen Geist überfordernde beeindruckende Anzahl von Möglichkeiten. Go hat aufgrund seiner Bauart aus einem Regelsatz von einigen wenigen simplen Regeln und Elementen jedoch die Konsequenz, dass es eine vielfach größere Anzahl von möglichen Figurenkombinationen in einem komplexen System zulässt (*in der Spieltheorie wird Go dem endlichen Nullsummenspiel mit perfekter Information zugeordnet*). Dieser Umfang an nicht formalisierbar auszudrückendem impliziten Wissen führte auch dazu, dass es mit den vorhandenen Rechenleistungen in einem „brute force“ Ansatz nicht möglich war, die Großmeister des Go-Spiels zu besiegen, in dem man mit einem „Durchprobieren“ aller möglichen Züge die impliziten mentalen Kapazitäten von menschlichen Großmeistern im Wettkampf schlagen konnte – durch die bloße Geschwindigkeit des Durchrechnens. „Alpha Go“ verwendete daher (2016) einen anderen Ansatz: tiefe neuronale Netzwerke und eine Baumsuche, wobei ein Regelnetzwerk trainiert wird mit einem überwachten bzw. bestärkendem Lernen und einem Bewertungsnetzwerk.

Inzwischen wurde diese Fähigkeit der Selbstmodifikation durch „Alpha Zero“ (2017) derartig gesteigert, dass nicht einmal Referenz-Partien eingegeben werden müssen. Das Programm lernt dadurch, dass es nur mit Kenntnis der Spielregeln solange gegen sich selber spielt, bis es alle Großmeister besiegen kann, die Jahrzehntlang intensiv gelernt und geübt haben, und die auf den Stufen von Jahrtausende-alten Hochkulturen des strategischen Denkens stehen. Die Lern- und Übungsphase des Programmes Alpha Zero dauert aber nur mehr ein paar Stunden. *Wir wissen aber nicht mehr, wie das Programm, diese selbstlernende KI, zu seinen Ergebnissen kommt.*

*Ironischerweise hat sich daher die Geschichte jetzt um 180 Grad gewendet:* Am Beispiel der traditionellen Strategiespiele Schach und Go hat sich seit etwa der Wende im Jahr 1989 gezeigt, dass zuerst die Rechner mit immenser Rechengeschwindigkeit die komplexen strategischen Problemlösungen besser simulieren konnten als unsere menschlichen Großmeister, die sich ihre Meisterschaft durch jahrzehntelange Bewährung in Turnieren errungen hatten. Danach bedurfte es seit 2015 der Rechengeschwindigkeiten im Stapeln von jenen neuronalen Netzen, die wir seit Jahrzehnten davor schon praktisch perfekt theoretisch modelliert haben, um mit selbstmodifizierenden Rechnern auch im Go-Spiel für Menschen unbesiegbar zu werden. Nunmehr versucht die US Verteidigungsforschungsagentur DARPA seit 2017 dahinter zu kommen, wie denn selbstmodifizierende Rechner zu ihren Ergebnissen gelangen. Sie verwendet dafür ein Konsortium von US Universitäten unter der Führung der State University Oregon – und ein Strategiespiel: „Star Craft“. *(Wer es nachlesen möchte: <http://scienceblog.sfu.ac.at/oregon-erklart-stanford-und-dem-mit-das-denken-von-kuenstlicher-intelligenz-bostons-roboter-marschieren-nach-tokio-und-black-pearl-sowie-fliegender-holl>).*

Man kann natürlich auch sagen „Dumm wie ein Sieb“ zur Künstlichen Intelligenz sagen - aber jede Wette: zu Zeiten Ingmar Bergmanns hätte es ganz anders geklungen...

## Beispiele aus der Forschung, Einsatz von Gehirnstromtechnik aktuell

### (Stand Anfang 2018)

# Ca. 31% der Vietnam-Veteranen, 10% aus dem Golf Krieg (Desert Storm), und 11% aus dem Krieg in Afghanistan leiden an PTSD. In einer Einzelstudie des Wake Forest Baptist Medical Center zeigen *Catherine L. Tegeler, Lee Gerdes, Hossam A. Shaltout, Jared F. Cook, Sean L. Simpson, Sung W. Lee, Charles H. Tegeler*, dass PTSD, hervorgerufen sowohl durch physische wie durch psychische Beeinträchtigungen, durch *HIRREM/High-resolution, relational, resonance-based, electroencephalic mirroring as a noninvasive, closed-loop, acoustic stimulation approach* nichtinvasiv behandelt werden kann. (Military Medical Research 2017; <https://doi.org/10.1186/s40779-017-0147-0>). Einschränkend wird angemerkt, dass diese Einzelstudie natürlich durch Doppelblind Versuche validiert werden müsste, da alle Teilnehmer in die Behandlung wissentlich eingebunden waren.

# 2017 erreichten *Chethan Pandarinath, Paul Nuyujukian, Christine H Blabe, Brittany L Sorice, Jad Saab, Francis R Willett, Leigh R Hochberg, Krishna V Shenoy, Jaimie M Henderson* ("High performance communication by people with paralysis using an intracortical brain-computer interface", vgl. <https://elifesciences.org/articles/18554>) den besten Wert von BCI-Kommunikation bei drei Patienten, von denen einer durch einen Unfall gelähmt wurde, und die beiden anderen an ALS/progressiver Paralyse leiden. Beim Buchstabieren wurde fast schon die Hälfte der Geschwindigkeit erreicht, die wir im Schnitt auf unseren Tastentelefonen erzielen.

# Grazer Forscherinnen und Forscher, *Andreas Pinegger, Hannah Hiebel, Selina C. Wriessnegger, Gernot R. Müller-Putz* schafften 2017 erstmals die Grundlage für Gehirnstrom-basiertes Komponieren: siebzehn musikinteressierten Laien schafften es mit Hilfe von EEG-Gehirn-Computer-Schnittstellen, immerhin über 90 Prozent der Töne vom konzentrierten Gedanken auf das Notenblatt zu übertragen. Bei dem zum Vergleich beteiligten Profi-Musiker kamen die erdachten Töne zu 98,2 Prozent korrekt auf das Notenblatt (*Composing only by thought: Novel application of the P300 brain-computer interface*; vgl. <http://journals.plos.org/plosone/article?id=10.1371/journal.pone.0181584> )

## Zusammenfassung

Wir stehen vor einer Flut von hochwirksamen Künstlichen Intelligenzen, deren Ergebnisse wir kaum vorhersehen können.

Diese unabsehbare Wirksamkeit baut sich auf durch die selbstmodifizierenden „deep learning“ Algorithmen, die auf die rasant wachsenden Datenmengen aller „Big Data“ greifen können, und die sich jetzt schon selbst reproduzieren können - also „KI/Künstliche Intelligenzen erschaffen durch KI/Künstliche Intelligenzen“.

Sie sind zwar (wahrscheinlich) meilenweit entfernt von auch nur etwas ähnlichem wie einem selbstreflexiven Bewusstsein.

Sie sind aber inzwischen schon fähig, unsere strategischen Kulturen und auch das implizite Können der erfahrensten Großmeister zu übertreffen, das auf Jahrtausende alten Zivilisationen im Osten wie im Westen aufgebaut ist, wie etwa im Schach und im Go-Spiel, wenn man ihnen nur die schlichten Regeln beibringt und ein paar Stunden Zeit einräumt, in der die KI/Künstliche Intelligenz übt, indem sie gegen sich selbst spielt. *Noch zu Zeiten des kalten Krieges hätte das wohl jede der Supermächte als ein unerreichbares Beispiel an kulturell konnotierter Intelligenz gewertet.*

Neben diesen selbstmodifizierenden „*deep learning*“ Algorithmen auf der Basis von gestapelten neuronalen Netzen werden in absehbarer Zeit neue Formen von neuronalen Netzen oder Clustern entstehen, die bei – weiterhin steigenden Rechenleistungen – unabsehbar wirkungsvollere Ergebnisse erreichen können.

Gleichzeitig tritt die Militärforschung nunmehr aus dem klinischen Experiment in den Bereich einer breit anwendbaren Gedankensteuerung, vor allem durch die Neurokognitionswissenschaften in den wichtigsten Prioritäten der Mensch/Maschine Schnittstellen (und die Militärmedizin in der Traumabehandlung von Gehirnschäden).

Dies reicht von einer neuroprothetisch gestützten Kommunikation durch „*silent talking*“ sogar mit Locked-In-Patienten, wie auch nicht-invasiv in beide Richtungen: etwa in der persönlichkeits-modifizierenden Behandlung durch (nichtinvasive) Gehirnstromtechnik bei schweren posttraumatischen Persönlichkeitsbeeinträchtigungen.

Dies wirft neben den klassischen ethischen Dilemmata die elementare Frage auf, wie lange die „Gedanken frei sind“, da sie etwa im „*silent talking*“ schon sehr früh auf- und abgegriffen werden können, bevor ein rechtlich zurechenbarer Ausdruck halbwegs zweifelsfrei ist.

Nun treffen diese ansteigenden Wellen aus den selbstmodifizierenden „*deep learning*“ Künstlichen Intelligenzen sowie aus der Verteidigungsforschung, insbesondere aus den Mensch/Maschine Schnittstellen, auf unsere „alten“ Rechtssysteme. Diese bieten zwar seit ca. 150 Jahren einen klassischen Grundrechtsschutz - wenn man aktuell unser seit 1867 geltendes Staatsgrundgesetz heranzieht, auch mit der ebenfalls erstmaligen Einsetzung eines Menschenrechtsgerichtshofes im Jahr 1867, und natürlich basierend auf den bürgerlich-revolutionären Märzartikeln von 1849.

Sie normieren zwar bis heute Glaubens- und Gewissensfreiheit, Gewerbe-, Vereins- und Versammlungsfreiheit, etc.etc., sowie die gültige Freiheit von Wissenschaft und Forschung sowie ihrer Lehre – *es gibt allerdings bisher kein gültiges Konzept von kognitiven und mentalen Freiheitsrechten über die aktuellsten Fragen der persönlichen Identität und Freiheit: wie lange gehört mir mein Gehirnstrom?*

Nun korrespondiert zwar die bisherige Entwicklung der (im internationalen Vergleich durchaus widersprüchlichen) Grundrechts-Normen zu den medizinischen Fragen an der Grenze des Lebens mit nahezu

allen unseren gesellschaftlichen und ökonomischen Institutionen (von der Organisation unserer Parlamente, der Gerichte, Universitäten, der modernen Kapitalgesellschaften bis zu unserem Haftungs- und Immaterialgüterrecht), deren heutige Ausformungen in etwa dieselbe historische Epoche fallen.

***Allerdings hat niemand auch nur im Traum daran gedacht hat, einen Gedankenaustausch auf Gehirnstrombasis in rechtliche Rahmen zu fassen, oder die Haftung zu normieren für eine hocheffektive aber „blinde“ KI/Künstliche Intelligenz, die ihrerseits erschaffen worden ist durch KI/Künstliche Intelligenzen, und es daher auch kein verbindliches Völkerrecht für den „Cyber-War“ gibt – jedoch mit einer Ausnahme: Pflichten und Rechte des Neutralen auch im „Cyber-War“ sind geregelt – in der Haager Landkriegsordnung.***