



Einsichten eines Wissenschaftsnarren (57)

KI: Kritik der schwätzenden Vernunft

Kann künstliche Intelligenz (KI) wirklich intelligent sein? Macht man sich klar, was Intelligenz überhaupt ausmacht, dann lautet die Antwort: Nein!

Künstliche Intelligenz (KI) und kein Ende: Täglich wird gewarnt, beruhigt, abgewägt. Sind ChatGPT *et al.* nun ein Segen für die Menschheit oder der Anfang von der Herrschaft der Maschinen? In der Berichterstattung werden dabei jede Menge Metaphern und Vergleiche benutzt, die die Leistungen von KI mittels Analogien vermenschlichen: Intelligenz, Lernen, Sprechen, Denken und Verstand, (Selbst-)Bewusstsein, Urteilen, Schließen, Entscheiden, Generalisieren, Fühlen, Kreativität, Irren, Halluzinieren, neuronale Netze und vieles mehr. Gleichzeitig werden Funktionen des menschlichen Gehirns mit Begriffen wie Computer, Memory, Speicher, Code, Algorithmus und so weiter beschrieben. Ebenso fehlt der Hinweis nicht, dass doch auch im menschlichen Gehirn elektrischer Strom fließt – ganz wie im Computer. Besonders das Feuilleton, benebelt von den mittlerweile verblüffenden Leistungen der schwätzenden und malenden Bots, treibt deshalb die Frage um, ob wir es bei der generativen KI nun schon mit „echter“ Intelligenz zu tun haben – gleichwohl die Frage im Namen doch eigentlich schon entschieden schien.

»Kann ein Computer mit Nullen und Einsen denken, kann er sich mittels KI zu einem geistigen Subjekt entwickeln?«

Warum gibt jetzt auch der Narr noch seinen Senf dazu? Weil er glaubt, dass die KI-Debatte voll am Thema vorbeigeht. Das Lager derer, die KI für intelligent halten, belegt dies mit einer Batterie von Leistungen, die alle ziemlich intelligent aussehen. Die Zweifler aber

überzeugt das nicht, ihnen fehlen immer noch bestimmte „Funktionalitäten“ von Intelligenz, die sie dann aus dem Ärmel ziehen – frei nach Teslers Theorem: „Intelligenz ist, was KI noch nicht gemacht hat.“ Die Diskussion bewegt sich bloß an der Oberfläche, statt sich mit der Frage zu befassen, was eigentlich Intelligenz, Denken, Sprache, Bewusstsein *et cetera* sind, um die KI daran zu messen.

Zum Glück hat sich vor über zweihundert Jahren schon mal jemand ganz grundsätzliche und sehr schlaue Gedanken zu eben diesen Geistestätigkeiten gemacht, welche die Debatte zurück auf eine inhaltliche Ebene führen können. Und das war nicht, wie der Titel dieser Zeilen vermuten ließe, Immanuel Kant, sondern sein Kritiker Georg Wilhelm Friedrich Hegel. Leider hat er seine Gedanken in eine für uns heutzutage schwer verdauliche Sprache verpackt, deshalb sind sie keine einfache Lektüre. Aber allemal gutes Material für den Narren, der im Folgenden versuchen wird, etwas zu leisten, was – Spoiler-Alarm! – KI eben nicht kann: Aus den Begriffen von Sprache und Denken abzuleiten, warum KI nicht sprechen und denken kann. Und zwar ganz prinzipiell nicht.

Die Kernfrage unserer Betrachtung lautet: Kann ein Computer mit Nullen und Einsen denken, kann er sich mittels KI zu einem geistigen Subjekt entwickeln – oder hat er das vielleicht sogar schon? Handelt es sich um einen „Verstand“, der möglicherweise drauf und dran ist, sich erkennend einen Begriff von der Welt zu machen? Und der dann anfängt, Gutes, aber vielleicht auch gar Schreckliches für und mit uns zu tun?

Beginnen wir auf der Ebene des Computers, genauer gesagt auf derjenigen der Transistoren. Ein Wort ist für den Computer – und damit für die KI – nichts als eine Folge von zwei physikalischen Zuständen: dem „Ein“ oder „Aus“ eines Schalters auf einem Halbleiter. Menschen, die die betreffenden Chips gebaut und programmiert haben, haben diesen beiden Zuständen Symbole zugewiesen: nämlich 0 und 1. Zahlen deshalb, weil man damit rechnen kann. Dies ist auch der Grund,

warum das Ding Computer heißt: Weil es, wie übrigens auch Ihr Handy, nichts anderes ist als eine programmierbare Rechenmaschine.

Der Programm-Code weist nun bestimmten Abfolgen dieser Symbole in vielen Zwischenschritten Worte zu, die nur für uns Menschen Bedeutung haben. Die Bezeichnung der dafür genutzten Algorithmen als „neuronale Netze“ ist nichts als ein gigantischer Marketing-Trick, genauso effektiv und falsch wie der Begriff künstliche „Intelligenz“ selbst. Tatsächlich sind künstliche neuronale Netzwerke nichts als mathematische Formeln, die mit begriffslosen Symbolen rechnen und durch

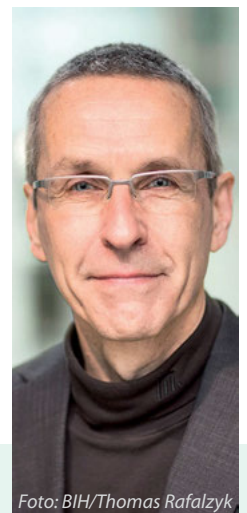


Foto: BIH/Thomas Rafalzyk

Ulrich Dirnagl

ist experimenteller Neurologe an der Berliner Charité und ist Gründungsdirektor des QUEST Center for Responsible Research am Berlin Institute of Health. Für seine Kolumne schlüpft er in die Rolle eines „Wissenschaftsnarren“ – um mit Lust und Laune dem Forschungsbetrieb so manche Nase zu drehen.

Sämtliche Folgen der „Einsichten eines Wissenschaftsnarren“ gibt es unter www.laborjournal.de/rubric/narr

sehr einfache und veraltete Modellvorstellungen von „echten“ Neuronen inspiriert wurden.

Hinter all dem, was da so gerechnet wird, kann folglich ganz prinzipiell keine Vorstellung oder ein Begriff der Sache stehen, die von der KI erfasst wurde. Auch wenn die KI noch so geschliffen argumentiert – es sind für sie inhaltsleere physikalische Zustände, codiert in Nullen und Einsen. Noch offensichtlicher wird das natürlich bei der Repräsentation von Bildern im Computer – auch deren Pixel sind für die KI nichts als binäre, gegenstandslose Schaltzustände von Transistoren.

»In der Bestimmung von mentalen Aktivitäten als „Fähigkeiten“ liegt die Erbsünde der fehlgeleiteten KI-Diskussionen.«

Damit ist eigentlich schon alles gesagt, warum KI nicht intelligent sein kann – und damit auch nicht sprechen, denken oder urteilen. Lassen Sie uns aber trotzdem noch ein bisschen weitermachen und den Fragen nachgehen, was Sprechen, Denken und Urteilen eigentlich sind. Dabei wird dann endgültig klar, warum das nicht durch Rechnen mit Nullen und Einsen geht.

Bei der Begriffsklärung dieser Geistesleistungen helfen leider keine der auch so populären, „funktionalen“ Bestimmungen weiter. Stellvertretend für die Schwäche solcher Definitionen sei hier diejenige von Intelligenz erwähnt, die die Konsensusgruppe führender internationaler Psychologen bevorzugt: „Intelligenz ist eine sehr allgemeine geistige Fähigkeit, die – unter anderem – die Fähigkeit zum schlussfolgernden Denken, zum Planen, zur Problemlösung, zum abstrakten Denken, zum Verständnis komplexer Ideen, zum schnellen Lernen und zum Lernen aus Erfahrungen umfasst.“ Das ist in Wahrheit keine Definition, sondern eine recht willkürliche Aufzählung von Fähigkeiten. Sie sagt nicht, was Intelligenz ist, sondern lediglich wozu man sie (möglicherweise) nutzen kann.

Ähnlich ist es bei der Definition von Sprache als „kommunizierende Verhaltensweise“. Oder wenn es über das Denken heißt: „Eine Form des Erkenntnisgewinns und der Erkenntnisnutzung; es ist etwas Dynamisches, das in der Zeit abläuft“.

Solche weithin durchgesetzten Un-Definitionen finden sich in Textbüchern der Psychologie oder auch bei einer Google-Suche. So kommen wir der Sache – also dem Begriff von Intelligenz, Sprechen und Denken – allerdings nicht näher. Denn genau in der Bestimmung von mentalen Aktivitäten als „Fähigkeiten“ liegt die Erbsünde der fehlgeleiteten KI-Diskussionen. Was ganz zwangsläufig dazu führt, dass man bei der geschmäckerischen Suche nach solchen Fähigkeiten in den Produkten der KI stecken bleiben muss.

Lassen wir zunächst jene zu Wort kommen, die glauben, dass der Rubikon nun endlich überschritten sei, und die Large Language Models (LLM) generelle und menschenähnliche Intelligenz besitzen. In Reinkultur findet sich die falsche Vorstellung vom intelligenten Computer in dem kürzlich veröffentlichten 155-seitigen Preprint „Sparks of Artificial General Intelligence: Early experiments with GPT-4“ (alle Links unter <http://dirnagl.com/lj>). Mit geradezu kindlicher Freude berichten die Wissenschaftler der Microsoft-Forschungsabteilung über ihre „Experimente“ mit einer Reihe von LLMs. Mit dabei natürlich GPT-4, der derzeitige Klassenprimus. Die Bots bekommen dabei Fragen und Aufgaben gestellt – und siehe da, die Resultate sehen doch ganz so aus, als ob die LLMs urteilen, empathisch und kreativ sind (schließlich malen sie und machen Musik!) sowie Selbstbewusstsein und „Theory of Mind“ besitzen. Natürlich attestieren die Forscher den LLMs noch einigen Verbesserungsbedarf: Manchmal, „halluzinieren“ sie, oder machen grobe Fehler, und das sogar bei simpelster Arithmetik. Ausgerechnet GPT-4 fällt beispielsweise in Mathe durch, weil es bei $7 \times 4 + 8 \times 8 = 88$ ausgibt. Aber eigentlich gilt den Autoren auch dies als Intelligenzbeweis nach dem Motto „Wie menschlich, wie allzu menschlich!“.

»Vielmehr haben sie genau das nicht geleistet, was eine der wesentlichen Leistungen menschlicher Intelligenz ist.«

Die Microsoft-Forscher berauschen sich förmlich an der sauberen Grammatik der LLMs und deren überaus höflichen Sprachstil, der mühelos zwischen Rap, Shakespeare und Py-

thon wechseln kann. Doch weil sie sich vor und während ihrer Spielereien keinen Begriff von ihrem (Forschungs-)Gegenstand gemacht haben, übersehen sie das Wesentliche. Das ist doppelt tragisch, denn die Autoren kommen deshalb nicht nur zu einem falschen Schluss (Computer = intelligent). Vielmehr haben sie darüber hinaus genau das *nicht* geleistet, was eine der wesentlichen Leistungen menschlicher Intelligenz ist – nämlich „sich einen Begriff von der Sache“ (hier also von der KI) zu machen. Schließlich macht man sich einen Begriff, indem man sich erkennend zur Welt stellt. Indem man also bestimmt, was die Sache *wirklich ist* – und nicht, wie sie einem vorkommt: Man benennt, was notwendig und wesentlich ist – und nicht, was nur zufällig und äußerlich ist.

»Die einzige Intelligenz, die hierbei im Spiel war, ist die menschliche, die die Software programmiert hat.«

Nur mit oberflächlichen Analogien argumentierend – eben Äußerlichkeiten statt Wesentlichem – liegen die Microsoftler deshalb voll daneben mit ihrem Schluss, dass sie es bei den LLMs mit allgemeiner oder irgendeiner anderen Form von Intelligenz zu tun haben. Sie erkennen nicht, dass die einzige Intelligenz, die da im Spiel war, die menschliche ist, die die Software programmiert hat – möglicherweise also ihre eigene! Wozu dann natürlich noch die geballte historische Intelligenz kommt, die für das Training verwendet wurde. Ebenfalls eine menschliche Intelligenz, die sich außerhalb und unabhängig von der KI betätigt hat – und so die Grundlage dafür geschaffen hat, dass die KI Erkennen, Verstehen und Entscheiden begriffslos simulieren kann, indem sie aus dem Material früherer Zuordnungen wiederum neue rein statistisch extrapoliert.

Der KI-Algorithmus stellt nämlich lediglich statistische Bezüge und Korrelationen zwischen Merkmalen der Eingabe her, egal ob diese aus Tweets von Elon Musk, Goethes Faust oder Wikipedia-Einträgen bestehen. Diese Bezüge zwischen den Inhalten des Trainingsmaterials sind rein stochastisch, sie beruhen nicht auf physikalischen, logischen oder inhaltlichen

Zusammenhängen. Entgegen anders lautender Behauptungen generalisieren die KI und ihr Sprachmodell dabei nicht, sondern schaffen bloß begriffslose Kennzeichnungen, Klassifizierungen und Regeln. Und diese beruhen eben nicht auf allgemeinen Bestimmungen, sondern sind lediglich das Resultat statistischer Ähnlichkeiten von Einzelfällen mit den Trainingsdaten.

»Das Resultat dieser Leistung der Intelligenz ist es, Sache und Namen im Denken eins werden zu lassen.«

Ein schönes Beispiel für diese Semantik-, Begriffs- und Inhaltslosigkeit der KI ist, dass sie Sprachen perfekt übersetzen kann, ohne die Vokabeln und die Grammatik von auch nur einer dieser Sprachen zu kennen oder zu verstehen – sie also sprechen zu können. Bei uns Menschen ist Letzteres aber die Grundvoraussetzung des Erlernens einer Fremdsprache. Entgegen landläufiger Meinung *lernt* die KI dabei auch nicht – es sei denn, man versteht wie viele Psychologen unter Lernen lediglich Konditionierung, Imitation oder Habituation. Nach dieser Definition ist Lernen jedoch lediglich stumpfsinniges Repetieren („Pauken“). Echtes Lernen bedeutet aber ein Erfassen des Lerngegenstandes durch Nachdenken oder Nachvollziehen – oder noch abstrakter und für KI unerreichbar: ein Erfassen der allgemeinen Bestimmungen einer Sache.

Das zeigt sich auch beim Spracherwerb. Ein Kind lernt nicht Sprechen durch das Abhören von Milliarden von Texten und nachfolgender statistischer Analyse. Es erlernt eine Sprache – und dabei gleichzeitig komplexes Denken, aber davon gleich! –, indem es aus eigener Erfahrung und Anschauung Vorstellungen im Gedächtnis „speichert“, und diese mit Sprachzeichen und Wörtern, die es hört, in eine feste Verbindung bringt. Das können demnach auch die Gebärden sein, die Taubstumme sehen und als ihre Sprache erlernen.

Dafür benötigt ein Kind erstaunlich wenig Material, auf jeden Fall keine Terabytes Weltliteratur. Das Gehirn des Kindes erlernt die Sprache durch deren Nutzung nach dem gehörten Vorbild und eignet sich deren grammatikalische Regeln an, ohne je eine Grammatik zu Rate zu ziehen. Das Resultat dieser Leistung der Intelligenz ist es, im Namen (zum Beispiel ein Wort oder Begriff) eine Sache zu erkennen und dabei beides – also Sache und Namen – im Denken eins werden zu lassen. Man muss sich keinen Baum mehr vorstellen, um beim Wort „Baum“ zu verstehen, was damit gemeint

ist – man könnte salopp auch sagen, das Wort „Baum“ ist im Gehirn zum Baum geworden. Wie das ein Gehirn mit einem synaptischen elektrochemischen Gewitter zustande bringt, ist gänzlich unbekannt. Aber wir müssen das auch gar nicht wissen, weil dieses neurobiologische Wissen nichts Zusätzliches beiträgt, es würde ja „nur“ die materiellen (physiologischen) Grundlagen des Denkens beschreiben – und nicht seinen Begriff, also was die Sache selbst ist.

In diesem „Embodiment“, dem Einswerden von Sache und Namen im Gehirn beim Denken, liegt auch der Grund, warum man mittels funktioneller Magnetresonanztomographie (fMRT) Hirnoxxygenierungsmuster „auslesen“ kann, die während dem Sprechen von Wörtern, dem Blick auf Bilder oder dem reinen Imaginieren von Bildern oder Gesprochenem auftreten. Diese Muster, die ihre Bedeutung im vorangegangenen Training mit eben diesen Bildern oder Worten zugewiesen bekommen haben, erlauben es dann, diese Wörter oder Bilder wieder teilweise zu rekonstruieren – allerdings nur im identischen, trainierten Individuum sowie auch nur mit hoher Fehlerrate. Das sind fantastische Ingenieurs- und Programmierleistungen, die auch für eine rudimentäre Kommunikation mit Gelähmten taugen könnte, die sich motorisch nicht mehr ausdrücken können (Brain-Computer-Interface), ist aber weder Gedankenlesen noch macht es die Maschine in irgendeiner Weise intelligenter: Der Computer findet inhaltsleere Muster, der Inhalt (= die Bedeutung) wird vom Menschen zugewiesen.

Durch die Sprache können wir über Dinge nachdenken, und dies auch ohne einen inneren Monolog zu führen. Das geht natürlich auch, manchmal mag es sogar hilfreich sein, insbesondere wenn wir komplexe Gedanken wälzen – wie gerade hier über Sprache und Denken. Doch auch ohne einen solchen Monolog beruht das Denken, mit dem wir die Welt verstehen lernen und mit dem wir dieses Wissen im täglichen Leben – wie auch gerade in der Wissenschaft – noch erweitern, auf Sprache.

»Wenn die KI dies dann ausgibt, vielleicht auch noch mit sonorer Stimme, scheint das manchem intelligent zu sein.«

Im Vergleich dazu die KI: Sie kann (oft) fehlerlos und umfassender als mancher Mensch „sagen“, was ein Baum ist. Aber sie spricht oder denkt dabei nicht. Denn die KI bedient sich nur beim Eintrag einer lexikalischen Liste von für

sie ansonsten inhaltslosen Bestimmungen, die sie für das Sprachzeichen „Baum“ aus Myriaden von Quellen zusammengesucht hat. Wenn die KI dies dann ausgibt, vielleicht auch noch mit sonorer Stimme, scheint das manchem intelligent zu sein. Aber die gleiche Person hält doch auch Wikipedia nicht für intelligent, nur weil man – wie auch die KI – in ihr beim Eintrag „Baum“ richtige Bestimmungen findet.

Es gibt also viele Gründe, warum eine KI nicht urteilen und Begriffe bilden kann. Diese liegen ganz grundsätzlich darin, dass die Welt in ihr in begriffsleeren Symbolen repräsentiert ist. Deshalb kann die KI auch nicht sprechen – und was uns als gesprochene Sprache verkauft wird, ist lediglich die Umsetzung von Zeichen in Töne. Weil aber schon das Zeichen für die KI keinen Inhalt hat, kann der daraus generierte Ton natürlich auch keinen haben. Und weil die KI nicht sprechen kann, kann sie auch nicht denken, denn die Sprache ist das Mittel des begrifflichen Denkens.

»Am Ende ist damit klar, dass KI keinen freien Willen entwickeln kann – und uns somit auch nicht an den Kragen gehen kann.«

Deshalb klappt es bei der KI dann auch nicht mit dem Urteilen, denn mit Sprache trennt die Intelligenz im Urteil das Subjekt von dessen Bestimmung, dem Prädikat (beispielsweise „Die Rose ist wohlriechend“, „Der Computer ist eine programmierbare Rechenmaschine“). Im Schluss beweisen wir daraufhin die Identität von Subjekt und Prädikat, womit man – wenn der Schluss richtig war – die Substanz einer Sache ausgemacht hat: Man hat sie erklärt, man hat sie unterschieden von dem, wie sie bloß erscheint oder vorkommt. Hegel würde sagen, man *hat* den Begriff der Sache, man *erfasst* die Realität im Gedanken. Wir Menschen können das – die KI nicht. Tiere übrigens auch nicht, weil sie zwar denken können, aber keine Sprache haben. Genügend Material für einen weiteren Artikel des Narren!

Am Ende ist damit ebenso klar, dass KI keinen freien Willen entwickeln kann – und uns somit auch nicht an den Kragen gehen kann, wie die KI „SkyNet“ im Film „Terminator“. Allerdings heißt das nicht, dass KI nicht gefährlich sein kann. Ihre schon länger genutzte Anwendung in der Militärtechnik beweist das ebenso wie die Fahrzeuge von Tesla, die im Autopilot-Modus manchmal ihre Eigner und dazu noch ein paar Fußgänger töten. Aber hier ist immer der Mensch das Subjekt, also der Gefährder. Ebenso wie bei Deep Fakes, Plagiaris-

mus und anderen kriminellen Aktivitäten, für die Menschen KI trefflich einsetzen.

Aus dem Bisherigen sollte daher klar geworden sein, dass sich alle bislang entwickelten KIs nur auf „next-word-“ oder „next-pixel-prediction“ verstehen – und damit kein neues Wissen schaffen können. KI schmeißt alles zusammen, was Menschen in eine digitale Form gebracht haben, vorausgesetzt dass es via Internet oder proprietäre Datenbanken verfügbar ist. Da findet die KI Richtiges und Nützliches, aber noch mehr Unsinniges, Unklares und Falsches. Damit repliziert KI natürlich auch alle existierenden Vorurteile. Weshalb Horden von Programmierern die resultierenden Unflätigkeiten, Volksverhetzungen, Gewaltaufrufe *et cetera* durch Zensur der KI wieder ausbügeln müssen. Oder gleich versuchen, diese Probleme durch Zensur auf der Ebene des kommunizierenden Menschen zu klären. Man verbietet der KI ganz einfach, zu antworten.

Mittels KI konfrontieren wir uns also mit den Leistungen und Auswüchsen unserer eigenen Intelligenz. Deshalb ist KI auch keine „künstliche Dummheit“, wie so mancher Kritiker glaubt. Auch deswegen, weil es für Dumm-

heit, die ja nichts anderes als der falsche Einsatz von Intelligenz ist, eine gute Portion Intelligenz braucht – und an der mangelt es der KI komplett.

»KI führt uns vor, wie geistlos viele unserer beruflichen Tätigkeiten sind. Und diese werden wohl tatsächlich durch KI ersetzt.«

KI taugt damit hervorragend zum Schreiben von Besinnungsaufsätzen und Gedichten. Und zu allem, bei dem menschliche Intelligenz auch nichts anderes macht als Muster zu erkennen, zu codieren, zu sortieren oder zu klassifizieren. Davon indes gibt es eine ganze Menge – in der Medizin, auf dem Amt, im Journalismus, beim Programmieren, dem Übersetzen oder auf dem Schlachtfeld. Da führt uns KI nur vor, wie geistlos doch viele unserer beruflichen Tätigkeiten letztlich sind. Und diese werden in nächster Zukunft wohl tatsächlich durch KI ersetzt.

Aber warum warnen eigentlich ausgerechnet die Vermarkter und Profiteure von KI so medienwirksam vor ihren eigenen Produkten? Und fordern Leute wie Elon Musk gar eine Trainingspause für ihre besten LLMs, „weil sie die Kontrolle über unsere Zivilisation übernehmen könnten“? Oder vergleichen sich selbst – wie Sam Altman, der Gründer von OpenAI – mit den „Vätern der Atombombe“? Zum einen wohl, weil sie selbst keinen Begriff von dem haben, was KI wirklich ist. Offenbar glauben sie tatsächlich, dass ihre LLMs generelle Intelligenz besitzen. Aber noch viel wichtiger: Sie präsentieren sich im Vorgriff auf staatliche Regulierung als verantwortungsbewusste Menschheitsbeglückter – und legen dabei den Turbogang im Hype um ihre Produkte ein.

Womit das Fazit klar wäre: An KI beunruhigt mich einzig die menschliche Intelligenz, die sie einsetzt, aber nicht die Aussicht, womöglich von Computern unterjocht zu werden.

Der Wissenschaftsnarr dankt Andreas Schneider für anregende Diskussionen. Weiterführende Literatur und Links finden sich wie immer unter: <http://dirnagl.com/lj>.



SIE SIND DOCH KEIN ROBOTER ...

BEFREIEN SIE SICH VOM ROUTINE-PIPETTIEREN

NEU: D-ONE
EINKANAL-PIPETTIERMODUL

INTEGRA



ASSIST PLUS automatisiert Pipetten

Automatisch durchgeführte Verdünnungsreihen, Reagenzzugaben, Probenumformatierungen, Normalisierungen und Hit-Picking sind damit für jedes Labor **erschwinglich**.



D-ONE – Einkanal-Pipettiermodul



VIAFLO - Elektronische Pipetten



VOYAGER - Pipetten mit einstellbarem Spitzenabstand

www.integra-biosciences.com