universitas

DAS MAGAZIN DER UNIVERSITÄT FREIBURG, SCHWEIZ | LE MAGAZINE DE L'UNIVERSITÉ DE FRIBOURG, SUISSE

03|202

Panta rhei 8

Albert Studer erzählt seinen Werdgang

Career Forum 50

Le 1er salon de recrutement de l'Unifr

Das Datum steht 56
Unifr und HEP | PH FR bald zusammen





Devenez acteur du savoir de demain Soutenez l'un de nos projets





Mehr erfahren und spenden Stipendien am Center For Transnational Legal Studies in London (CTLS)

Ideale Vorbereitung für eine internationale Karriere

Programme d'enseignement sur l'histoire des entreprises et des industries du XIX^e siècle à nos jours

Forger l'avenir par la connaissance du passé



En savoir plus et faire un don

Grâce à votre don, vous permettez à nos étudiant·e·s de continuer à bénéficier de ces possibilités uniques et enrichissantes. Merci de tout cœur!

Mit Ihrer Spende ermöglichen Sie unseren Studierenden, weiterhin von diesen einzigartigen und bereichernden Möglichkeiten zu profitieren. Ganz herzlichen Dank!

Impressum

universitas

Das Wissenschaftsmagazin der Universität Freiburg Le magazine scientifique de l'Université de Fribourg

Herausgeberin | Editrice

Universität Freiburg Unicom Kommunikation & Medien www.unifr.ch/unicom

Chefredaktion | Rédaction en chef

Claudia Brülhart | claudia.bruelhart@unifr.ch Farida Khali (Stv./adj.) | farida.khali@unifr.ch

Art Direction

Daniel Wynistorf | daniel.wynistorf@unifr.ch

Adresse

Universität Freiburg Unicom Kommunikation & Medien Avenue de l'Europe 20, 1700 Freiburg www.unifr.ch

Online | En ligne

www.unifr.ch/universitas

Autor_innen | Auteur·e·s

Jean-Luc Brülhart | jeanluc.bruelhart@yahoo.com Christian Doninelli | christian.doninelli@unifr.ch Matthias Fasel | matthiasfasel@hotmail.com Roland Fischer | roland@scifischer.net Pierre Jenny | pierrejenny@bluewin.ch Patricia Michaud | info@patricia-michaud.ch Ori Schipper | ori _schipper@sunrise.ch

Titelbild | Image de couverture

Roland Fischer | roland@scifischer.com stablediffusionweb.com

Illustrationen Dossier | Illustrations du dossier

Roland Fischer | roland@scifischer.com stablediffusionweb.com

Fotos | Photos

Stéphane Schmutz | info@stemutz.com Thomas Delley | info@thomasdelley.com Getty Images | gettyimages.com

Sekretariat | Secrétariat

Marie-Claude Clément | marie-claude.clement@unifr.ch

Druck | Impression

Canisius SA Avenue Beauregard 3 CH-1700 Fribourg

Auflage | Tirage

9'400 Exemplare | dreimal jährlich 9'400 exemplaires | trois fois par année

ISSN 1663 8026

Alle Rechte vorbehalten.

Nachdruck nur mit Genehmigung der Redaktion.

Tous droits réservés.

La réimpression n'est autorisée qu'avec l'accord de la rédaction.

Die nächste Ausgabe erscheint im April 2024.

La prochaine édition paraîtra en avril 2024

Die in dieser Publikation zum Ausdruck gebrachten Meinungen entsprechen nicht zwangsläufig der Haltung der Unifr.

Les opinions exprimées dans les articles d'universitas ne reflètent pas forcément celles de la rédaction.



UNIVERSITÉ DE FRIBOURG

Edito

Ich habe mich entschieden – und werde dieses Edito *nicht* von ChatGPT schreiben lassen. Nur: Wie kann ich das beweisen? Gar nicht. Sie müssen mir glauben, sich auf ihr Gespür verlassen, auf den vielzitierten «gesunden Menschenverstand». Ginge es nach mir, wäre dieser Ausdruck «Wort des Jahres 2023». Aber es geht nicht nach mir (das Wort des Jahres 2023 ist «Monsterbank», gewählt von Mensch und Maschine).

Ausserdem gibt es den gesunden Menschenverstand ja schon viel zu lange - handelt es sich beim Wort des Jahres doch häufig um neuere Kreationen wie etwa Impfdurchbruch oder Strommangellage. Tatsächlich hat der gesunde Menschenverstand im deutschen Sprachgebrauch während der Zeit der Aufklärung an Bedeutung gewonnen und sich diskret etabliert. Und jetzt: Bumm! Der gesunde Menschenverstand feiert ein wahres Comeback. Er gilt sozusagen als Benchmark im Umgang mit der künstlichen Intelligenz. Sagt unser Verstand, der «gesunde» wohlverstanden, dass die KI-Entwicklungen und -Auswüchse ethisch noch zu verantworten sind, machen wir weiter. Und sollte dieser Wächter über Gut und Böse mal anschlagen... dann was? Diese Frage kann kein Mensch und keine Maschine beantworten. Nur der gesunde Menschenverstand, der hat vielleicht eine leise Ahnung.

Ein Wort noch zu den Illustrationen im Dossier dieser Ausgabe: Die Bilder wurden, wie Sie vielleicht schon geahnt haben, von KI generiert. Kurz zusammengefasst heisst das: Wir haben ChatGPT die Artikel gefüttert und um Stichworte gebeten, die zu einer Illustration des Artikels führen sollten. Diese Stichworte spiesen wir in Form von Prompts dem Bildgenerator stablediffusionweb.com mit Angaben zum Stil, den wir gerne hätten. Das Resultat ist verblüffend. Es ist toll. Nur beim genaueren Hinschauen sieht man: Es ist nicht perfekt. Ist das nun beruhigend?

Ein Hoch auf den Menschenverstand im 2024 – und frohe Festtage!

Herzlich, Claudia Brülhart Chefredaktorin

Inhalt | Sommaire

News

← Habemus doctores

Die Universität Freiburg zeichnet fünf Persönlichkeiten mit der Ehrendoktorwürde aus

Porträt

8 Keine klassische Karriere

Albert Studer hat als Berufsberater für den Kanton Freiburg vielen den richtigen Weg gewiesen. Über seine eigene Laufbahn muss er selber schmunzeln

- 10 Dossier Künstliche Intelligenz
- 12 Lauter Gretchenfragen
 Im Zentrum: Der gesunde Menschenverstand
- Du phénomène technologique à l'objet social Poser un regard nuancé sur l'empreinte sociétale de l'IA
- 20 The Singularity is Near! Wird Fiktion bald Realität?
- 23 **Label für digitale Nachhaltigkeit**Ethisch denken ist nicht schwer ethisch handeln umso mehr
- 25 **L'IA à l'école ou l'école de l'IA**Pour une relation sereine entre éducation, formation et technologies émergentes
- 28 Un contrat? Adressez-vous à mon assistant IA!
 Contrats algorithmiques: comment protéger les
 consommatrices et les consommateurs?
- 31 **Asymmetrie zwischen Lob und Tadel**Wenn eine KI die Hausaufgaben macht
- La conscience artificielle entre Pygmalion et Pandore
 Corps, âme, esprit: l'être humain bionique est-il envisageable?
- 38 **«Pics or it didn't happen» war gestern**Können wir unseren Augen und Ohren noch trauen?









Forschung & Lehre

42 Der Feind meines Feindes

Invasive Arten sind weltweit ein immer grösser werdendes Problem. Eine Lösung sind in gewissen Fällen biologische Agenten – wie etwa der Rüsselkäfer. Der Biologe Sven Bacher erklärt, was es damit auf sich hat

Recherche & Enseignement

46 Révolution copernicienne chez les neuroscientifiques

Le cerveau n'est pas seul au centre de la perception visuelle

Interview

50 Un pont entre deux mondes

Le Career Forum veut rapprocher milieux académiques et économiques fribourgeois. Entretien avec Jerry Krattiger, directeur de la Promotion économique du Canton de Fribourg, et Astrid Epiney, rectrice de l'Université de Fribourg

Fokus

56 Uni und PH gründen eine neue Fakultät

Ab August 2025 werden die Universität Freiburg und die Pädagogische Hochschule gemeinsame Sache machen: Institutionell vereint in einer gemeinsamen Fakultät

People & News

F7 Prix et nominations

Quoi de neuf à l'Unifr?

Red & Antwort

58 Wolfgang Taube

Professor für Sportwissenschaften



Habemus doctores



Am 15. November hat die Universität Freiburg Ehrendoktorwürden an folgende Personen verliehen: Luca Vetterli, Geschäftsführer von Pro Natura Ticino, für sein Engagement für den Moor- und Gewässerschutz zu Gunsten der Biodiversität in der Schweiz; Ann Blair, Geschichtsprofessorin an der Harvard University, für ihre Forschung zur Beziehung zwischen Wissenschaft, Religion und Politik; Mark Musen, Professor für Biomedizinische Informatik an der Stanford University, der mit Protégé die weltweit führende Software zur Erstellung und Verwaltung von Terminologien und Ontologien entwickelte; Pascal Ancel, Professor für französisches Recht, für sein Werk über den Vergleich zwischen französischem, belgischem und luxemburgischem Recht sowie zur fundierten Analyse der Auswirkungen eines mehrsprachigen Systems wie das des luxemburgischen Rechts sowie Physiker, Philosoph und Theologe Hans-Dieter Mutschler für seine interdisziplinäre Kompetenz im Gespräch zwischen Theologie und Naturwissenschaften.

unifr.ch/go/dies



Von links nach rechts, sitzend: Luca Vetterli, Ann Blair, Mark Musen, Pascal Ancel und Hans-Dieter Mutschler
Von links nach rechts, stehend: Die Dekane Ulrich Ultes-Nitsche, Math-Naturwissenschaftliche und Medizinische Fakultät;
Dominik Schöbi, Philosophische Fakultät; Dirk Morschett, Wirtschafts- und Sozialwissenschaftliche Fakultät;
Jacques Dubey, Rechtswissenschaftliche Fakultät; Joachim Negel, Theologische Fakultät

Keine klassische Karriere

Wenn es um Bildung geht, macht ihm niemand was vor: Albert Studer war 13 Jahre lang als Berufsberater tätig und hat auch selber einen Parcours, der einen guten Teil der Schweizer Bildungslandschaft umfasst. Claudia Brülhart

Albert Studer, Sie haben an der Uni Freiburg studiert – aber nicht auf direktem Weg, wenn man so sagen kann.

Man kann. Nach der obligatorischen Schulzeit begann ich zuerst die Ausbildung am Lehrerseminar. Ich musste allerdings nach einem knappen Jahr abbrechen, Klavier und Blockflöte lagen mir ganz und gar nicht. Mangels besserer Ideen habe ich dann eine Schreinerlehre absolviert – und danach am Collège die Matura nachgeholt. Ich brauchte Zeit zum Nachdenken.

Das Nachdenken hat Sie an die Unifr geführt, zum Studium der Heilpädagogik.

Nicht das Nachdenken, nein (*lacht*). Ich habe nach der Matura für die Stiftung «Les Buissonnets» gearbeitet und dort meine künftige Frau getroffen, eine Praktikantin, die an der Uni Heil- und Sonderpädagogik studieren wollte. Also habe ich mich auch eingeschrieben. Hinzu kam, dass die Universität für mich die einzige Möglichkeit war, um studieren zu können, da ich nebenbei 80 Prozent arbeiten musste.

Nach einem Jahr Arbeit als Heilpädagoge zog es Sie in die Jugendarbeit. Später leiteten Sie den Verein für aktive Arbeitsmarktmassnahmen. Dafür haben Sie erneut die Schulbank gedrückt und einen CAS in BWL für Führungskräfte gemacht.

Die Jugendarbeit hat mir sehr gut gefallen – sie hat mich aber auch beschäftigt. Ich habe viele Jugendliche kennengelernt, die Schwierigkeiten hatten, ihren Weg zu finden. Als ich dann erfahren habe, dass mein ehemaliger Berufsberater in Pension geht, hat der Funken wieder gezündet.

Sie wurden Berufs- und Laufbahnberater beim Kanton Freiburg.

Ich hatte Glück, diese Stelle zu kriegen. Ich war ja nicht ausgebildeter Berufsberater. Mit meiner Erfahrung in der Jugendarbeit und dem Diplom als Heilpädagoge konnte ich den Kanton überzeugen. Ich habe die Stelle angetreten und einen Master gemacht in Berufs-, Studien- und Laufbahnberatung an der ZHAW.

War dieser Wechsel ein Kopf- oder ein Bauchentscheid?

Es war beides. Mit meinen mittlerweile 42 Jahren und als mehrfacher Vater wusste ich, dass ich diesen Jugendlichen etwas mit auf den Weg geben konnte – nicht zuletzt durch meine eigenen Erfahrungen. Ich war damals ein grosser Fan des dualen Systems, für das die Schweiz ja bekannt ist. Also eine Lehre zu absolvieren mit dem Besuch der Berufsschule. Und danach, bei Bedarf, via Berufsmatura und Passerelle an eine Hochschule oder Fachhochschule.

Ihre Formulierung lässt schliessen, dass Sie heute nicht mehr ganz so überzeugt sind von diesem dualen Bildungssystem?

Das duale System mit dem Verkaufsargument «Lerne Coiffeuse, werde Biologin» – einst ein Werbespruch für dieses System – macht keinen Sinn. Wir klagen über Fachkräftemangel und sagen, dass zu viele die Matura machen würden. Aber das stimmt nicht. Heute gibt es die Berufsmatura, die man während und nach der Lehre machen kann – und DAS machen viele. Und so finden wir Lehrabgänger_innen dann an der Uni oder an einer Fachhochschule, wo sie Physik studieren oder Lehrer_innen werden.

Sie sind auch ein Beispiel für dieses System: Lerne Schreiner – werde Berufsberater

Aber das System ist nicht mehr zeitgemäss. Wir sollten die Lehre an die heutige Gesellschaft, die heutigen Jugendlichen anpassen. Ein gutes Beispiel dafür ist die «Ecole du Métal» für angehende Metallbauer_innen. Das erste Jahr der Ausbildung verbringen sie in der Schule mit Peers – das macht den Übergang von der obligatorischen Schulzeit in die Lehre viel einfacher.

Haben Universitäten an Attraktivität verloren unter den Jugendlichen?

Nein, viele möchten an die Uni – aber nicht alle erfüllen die Anforderungen am Ende der obligatorischen Schulzeit. Also versuchen sie es über die Lehre und die Berufsmatura mit anschliessender Passerelle.

Ist es durch die vielen Möglichkeiten nicht einfacher geworden für die Jugendlichen?

Jein. Heute stehen 240 Berufe zur Auswahl! Aber die Jugendlichen müssen sich zu früh entscheiden. Eine europäische Vergleichsstudie zeigt, dass Europäerinnen zwischen 17 und 19 die Berufswahl treffen, die unsere Mädchen hier zwischen 13 und 15 treffen sollten. Das Resultat: Mädchen wählen mehrheitlich unter 12 Berufen aus – von 240! Dabei weiss man: Bei den Mädchen stabilisiert sich die pubertäre Phase so mit 17, 18 Jahren – sie werden mutiger, entscheidungsfreudiger. Bei den Jungs dauert die Pubertät ja mitunter bis 50 (*lacht*).

Seit kurzem sind Sie Sektionschef bei der kantonalen IV-Stelle und leiten die Eingliederung von Jugendlichen ins Berufsleben. Diese Stelle vereint Vieles, das in Ihrem Curriculum vorkommt.

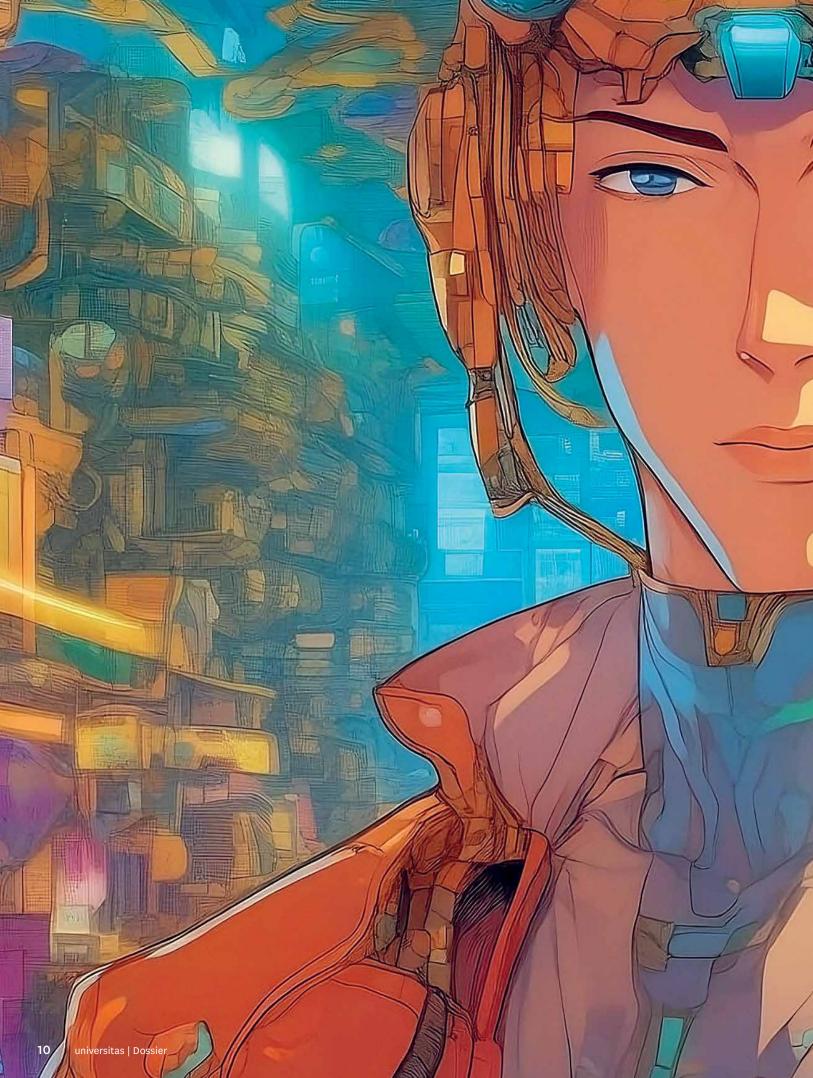
Ich möchte keine meiner Erfahrungen missen und denke, dass ich damit gut aufgestellt bin für diese neue Herausforderung. Ich sehe den Sinn der Massnahmen, die wir bei der IV anwenden – seit Anfang 2022 gibt es neu die Frühintervention ab 13 Jahren – und ich kann abschätzen, wann welche Unterstützung Sinn macht.

Und der Schreiner in dem Ganzen?

Gerade heute hat mir jemand gesagt er sei froh, dass ich einen handwerklichen Background habe – und nicht nur einen Universitätsabschluss in der Tasche.

Claudia Brülhart ist Chefredaktorin des Wissenschaftsmagazins «universitas».







Lauter Gretchenfragen

Kann eine Künstliche Intelligenz (KI) gut oder böse sein? Wie weit kann man einer KI gesunden Menschenverstand beibringen? Und überhaupt: Wer weiss das schon? Expert_innenrunde mit Juristin Nula Frei und den Informatikexperten Denis Lalanne und Simon Ruffieux. Roland Fischer

Eine kursorische Frage, zum Einstieg: Wie baut man eine KI, die sich ethisch verhält? Oder: Wie baut man Ethik in eine KI ein?

Nula Frei: Ich weiss gar nicht, ob ich die Kompetenzen habe, darauf zu antworten. Ich weiss weniger wie man eine KI bauen sollte, als vielmehr, wie wir sie regulieren sollten. Denis Lalanne: Wenn wir von Kompetenzen reden: Das mag nun vielleicht ein wenig provokativ klingen, aber ich habe den Eindruck, dass es in vielen Bereichen ausserhalb der Computerwissenschaften an KI-Fachwissen fehlt. Wie reguliert man also etwas, das man gar nicht recht versteht? Die Modelle werden immer komplexer, selbst die KI-Spezialist_innen haben zuweilen Mühe zu erklären, wie sie genau funktionieren.

Nula Frei: Das ist natürlich die grosse Frage. Wir haben es auf jeden Fall mit einem «moving target» zu tun, die Technologie entwickelt sich ja so schnell. Das erzeugt natürlich Reibung zwischen Entscheidungsträgern und Rechtsexperten auf der einen und Forschenden auf der anderen Seite. Es geht letztlich um die grosse Frage, wie viel technische Expertise vorhanden sein muss auf Seiten der Politik. Ich würde mal sagen: Der interdisziplinäre Austausch wird auf jeden Fall immer wichtiger.

Denis Lalanne: Das sehe ich auch so, Interdisziplinarität ist der Schlüssel. Eine andere Option gibt es gar nicht. Dabei ist mir aufgefallen, dass es gerade die Ethiker sind, die häufig fehlen in den entsprechenden Kommissionen.

Es war tatsächlich sehr schwierig, eine Ethik-Fachperson zu finden, die sich auch mit KI auskennt. Und das gilt nicht nur für die Unifr.

Denis Lalanne: Dabei gäbe es genug Problemstellungen, manche haben schon fast Klassikerstatus. Kennen Sie zum Beispiel das Trolley-Problem?

Sie meinen das Gedankenexperiment, das unseren Umgang mit utilitaristischer Ethik aufzeigt? Ein Tramwagen wird unweigerlich einen Unfall verursachen und man muss entscheiden, wie die Weiche gestellt werden soll.

Denis Lalanne: Genau. Im Zusammenhang mit selbstfahrenden Autos wird das plötzlich sehr konkret. Unlängst gab es am MIT eine Studie, die global Daten dazu gesammelt hat und deutliche interkulturelle Unterschiede bei den ethischen Präferenzen ergab. Da frage ich mich: Muss auch die Rechtsprechung entsprechend lokal ausgeformt werden? Nula Frei: Natürlich muss Recht grundsätzlich den ethischen Standards einer jeweiligen Gesellschaft entsprechen. Das heisst aber nicht, dass wir kleinteilige und untereinander inkompatible Regelungen anstreben, im Gegenteil. Meiner Ansicht nach ist gerade KI ein Paradebeispiel dafür, dass wir übergreifende Regelungen brauchen. Wir haben es hier mit allgemeinen Herausforderungen zu tun, die – zumindest in Europa – auch einen gemeinsamen gesetzgeberischen Rahmen brauchen.

Denis Lalanne: Ich bin allerdings sehr froh, dass sich die Regulierungsbemühungen in Europa von denen in den USA unterscheiden. So wie dort derzeit reguliert wird, würde ich ganz ehrlich nicht mit KI-Anwendungen leben wollen.

Simon Ruffieux: Oder in China!

Nula Frei: Da gibt es auf jeden Fall grosse Unterschiede. Insbesondere wenn wir uns ansehen, welche Anwendungen nicht reguliert, sondern rundweg verboten werden sollten: Das unterscheidet sich sehr wesentlich zwischen Europa, Amerika und China.

In welchen Feldern zum Beispiel?

Nula Frei: Es scheint in unserem kulturellen Kontext zum Beispiel einen Konsens zu geben, dass wir nichts wollen, was einem Social Scoring nahekommt, wie es gerade in China Realität wird. Gesichtserkennung in Echtzeit im öffentlichen Raum dagegen: Das wird gerade sehr lebhaft debattiert, das scheint mir noch nicht entschieden. Insofern ist es letztlich eine Frage der demokratischen Entscheidung, welche Nutzungen der KI wir wollen und wo wir Grenzen ziehen.

Simon Ruffieux: Ich bin nicht sicher, ob das funktionieren wird. Eines der Probleme mit Ethik und KI ist, dass die Implementierung und Nutzung schwer zu kontrollieren ist. Wenn ich etwas in einem laschen regulatorischen Rahmen entwickeln kann, dann kann es potentiell überall in der Welt zum Einsatz kommen. Gerade China ist da derzeit sehr offen, da werden KI-Anwendungen entwickelt, die potentiell problematisch sind. Insofern sehe ich auch Open Source kritisch – können wir, wenn die Modelle ganz offen sind, noch kontrollieren, wer was wie nutzt? Ich sehe da tatsächlich ein Missbrauchspotential von sehr mächtigen KI-Werkzeugen, die für jedermann leicht zugänglich sind.

Aber gleichzeitig gibt es die Forderung, KI-Modelle möglichst transparent zu machen, das würde doch für Open Source sprechen?

Simon Ruffieux: Ja. Ich sehe auch, wie wichtig es ist, die Modelle transparenter zu machen und dafür zu sorgen, dass möglichst alle Zugang zu diesen Technologien bekommen, dass also die Entwicklung nicht hinter verschlossenen Türen passiert. Aus dieser Perspektive ist die Demokratisierung der KI natürlich sehr zu begrüssen. Es ist ein komplexes Problem – ohne simple Lösung.

Was würde das denn heissen, auf einer ethischen Ebene? Über Pro und Kontra demokratisch zu entscheiden?

Denis Lalanne: Da passiert ja gerade einiges, es entstehen eine ganze Reihe von Initiativen, die meisten mit einem eher technischen Fokus. Was mir fehlt ist eine Koordination dieser Initiativen, um dieses Wissen, diese Reflexion zu den Entscheidungsträgern bringen.

Nula Frei: Wobei sich natürlich die Frage stellt, wann der Staat überhaupt intervenieren muss. Eigentlich sagt man: Sobald es Auswirkungen gibt, die Menschen direkt schaden. Typischerweise ist das der Fall, wenn es ein grosses Machtgefälle gibt, wenn Dinge zur Anwendung kommen, die für die meisten Menschen nicht zu durchschauen und zu verstehen sind oder gegen die man sich nicht wehren kann. Dann werden wir als Nutzerinnen und Nutzer verwundbar. Denis Lalanne: Aber wann genau kommt denn diese Schwelle? Was wäre ein «Regulations-Trigger»? Es ist ja so, dass die Entwicklungen kontinuierlich vorangehen, in kleinen Schritten. Das erinnert ein wenig an die Art und Weise, wie Social Media zur gesellschaftlichen Normalität wurde: Zuckerberg und Konsorten trieben das bestimmt ohne bösen Willen voran, aber es wurde immer bedeutender und allmählich zeigten sich auch gefährliche Effekte.

Nula Frei: Das ist auf jeden Fall ein Problem. Leider fehlt uns oft das Verständnis dessen, was gerade passiert. Sei es auf gesetzgeberischer Seite oder im Parlament – da wurden bei den letzten Wahlen ja sogar Digital-Expert_innen aus dem Nationalrat wieder abgewählt. Es ist wirklich eine Herausforderung, auf dem Laufenden zu bleiben und nicht einfach von den Entwicklungen überrollt zu werden. Sonst geht es uns wie dem Frosch im sich langsam aufheizenden Wasser, der am Schluss bei lebendigem Leibe gekocht wird, weil er der graduellen Veränderung der Temperatur keine Beachtung schenkt.

Und gerade als Politiker_in will man dem Fortschritt ja nicht im Wege stehen.

Nula Frei: Ja, wir folgen da oft auch ein wenig blind diesem Fortschritts-Paradigma, das hat gerade das Beispiel Social Media schön gezeigt. Es stimmt, Facebook hatte sicher keine bösen Absichten, dafür spielten umso stärkere kommerzielle Anreize. Da bedient man sich natürlich gern dieser positiven Fortschrittserzählung, um sein Geschäftsmodell zu legitimieren. Mitunter beschleicht mich das Gefühl, digitale Tools würden eingesetzt, weil sie für «Fortschritt» stehen und nicht weil sie einen echten Nutzen bieten.

Simon Ruffieux: Unbedingt. Und oft wissen wir gar nicht, wo überall sie eingesetzt werden, gerade im Fall der KI. Auch da wieder: Man nutzt KI-Tools ganz einfach deshalb, weil sie zur Verfügung stehen, ohne wirklich zu prüfen ob sie auch wirklich leisten können, was man sich von ihnen verspricht. Vor allem bei Behörden ist das beunruhigend. Zum Beispiel bei der Zürcher Polizei, die Tools nutzt, um die Rückfallquote von Straffälligen abzuschätzen. Dabei wurde von der Forschung wiederholt aufgezeigt, dass diese KI-Technologien nicht ausgereift sind, dass sie Vorurteile reproduzieren und auch gerne falsche Vorhersagen liefern. Trotzdem kommen sie zum Einsatz, ohne dass man sich die Zeit und Mühe nehmen würde, die Nutzer_innen über die Fallstricke aufzuklären. Das ist der Umstand, der mir derzeit am meisten Sorge macht.

Denis Lalanne: Ich bin da ehrlich gesagt hin- und hergerissen. Ich gebe Simon recht in seiner Kritik, aber man muss auch sagen: Die KI hilft uns eben auch, solche Biases überhaupt zu sehen. Insofern kann dieselbe Technik ein Mittel zur Diskriminierung sein und ein Mittel, sich solcher Diskriminierungen bewusst zu werden. Ich finde es insofern immer schwieriger, in Kategorien von «guter» und «böser» KI zu denken.

Dasselbe sagt man auch von einem Auto, oder von einem Kernkraftwerk: Es kommt auf die Nutzung an, und Risiken gehören inhärent zu Technologie. Insofern könnten wir auch auf ältere Technikethik zurückgreifen. Könnte man diese Frameworks nicht auch auf KI anwenden?

Denis Lalanne: Ja und Nein. Ja, es stimmt, wir mussten uns immer wieder mit radikal neuen Technologien auseinandersetzen. Insofern können wir bestimmt bestehende Rahmenbedingungen auch auf KI anwenden, zum Beispiel um sicherzustellen, dass wir keine KI entwickeln, die Diskriminierung befördert. Das kategorisch Neue liegt meiner Ansicht nach in der Geschwindigkeit der Veränderung, hier liegt die grosse Schwierigkeit.

Wobei: Über Diskriminierung mussten wir ja noch nie nachdenken im Zusammenhang mit Technologie – wenn ich ein Auto baue, zum Beispiel, dann brauche ich mir darüber nicht den Kopf zu zerbrechen.

Nula Frei: Es ist allerdings eine gute Frage, inwiefern wir auf frühere Rahmenbedingungen zurückgreifen können. Im Zusammenhang mit Diskriminierung kommt mir höchstens Forschung mit genetischen Daten in den Sinn, da gab es meines Wissens eine entsprechende regulatorische Diskussion. Ich glaube tatsächlich, dass KI neue regulatorische Herausforderungen darstellt, weil es nicht einfach gebaut und importiert wird, es ist transnational, es wird in verschiedenen Kontexten eingesetzt, ob es nun Open Source ist oder nicht. Wir können also nicht einfach alte Regeln der Produkthaftung übernehmen, die sich bei technischen Geräten bis jetzt bewährt haben.

Denis Lalanne: Ja, da verschiebt sich etwas. Bisher war Technik einfach ein Werkzeug, jetzt geht es eher in die Richtung, menschliche Intelligenz nachzubauen, wenn nicht gar zu ersetzen. KI kann autonom entscheiden – da müssen wir uns klar machen, in welchen Kontexten das akzeptabel und in welchen es problematisch ist. Am Flughafen, wenn ein automatisiertes Sicherheitssystem Waffen im Gepäck erkennt: Das ist wohl unproblematisch. Aber in Gerichtsprozessen, wo es ein gutes Mass an «gesundem Menschenverstand» braucht?

Nula Frei: Ich habe eben von einem System gelesen, das Kamerabilder von Tramstationen in Zürich auswerten kann: Gibt es Rauch, liegt jemand am Boden und braucht Hilfe? Dann kann entsprechend automatisch Alarm ausgelöst werden – so etwas wäre ja sehr zu begrüssen. Aber wie sieht es bei der Ärztin aus, die mit einem KI-Assistenten arbeitet? Muss sie genau verstehen, wie diese KI funktioniert, um sie im ärztlichen Alltag einsetzen zu können?

In welcher neuen Kategorie könnten wir denn denken, wenn wir es nicht mehr einfach mit «Werkzeugen» zu tun haben? Sollten wir solche Assistenz-Systeme eher als «Kollegen» betrachten, denen wir ethische Frameworks nicht vorschreiben, sondern beibringen?

Simon Ruffieux: Kollegen? Das hätte übrigens ganz entscheidend auch mit der Haftungsfrage zu tun. Der Kollege im Spital sollte immer in der Lage sein zu erklären, wie er zu dieser oder jener Einschätzung gekommen ist. Und in dem Moment übernimmt er auch zumindest einen Teil der Verantwortung, oder rechtlich gesehen: der Haftung für eine Fehlentscheidung. Wir müssen uns klar machen: Auch die allerbesten KI-Systeme werden nicht ganz fehlerfrei arbeiten. Also sagen wir mal, ein Krebs-Assistenzsystem liegt in 99 Prozent der Fälle richtig – möglicherweise sogar öfter als es ein menschlicher Arzt tut. Aber das heisst eben auch: es wird 1 Prozent Fälle geben, wo Leute sterben, weil das System falsch entschieden hat. Wenn wir diese KI als «Kollegen» ansehen, müsste sie uns erklären können, wo der Fehler passiert ist.

Und müsste sie sich vielleicht sogar schuldig fühlen? Es ist ja irgendwie unvorstellbar, dass sich eine KI je einem Gerichtsprozess stellen müsste, oder?

Nula Frei: Unter aktuellem Recht ist das tatsächlich ausgeschlossen. Es muss immer einen Menschen geben, der verantwortlich ist für die Nutzung der KI. Aber eben, im oben genannten Fall: Haftet dann automatisch die Ärztin? Diese Fragen sind von den Gerichten noch nicht abschliessend beantwortet worden, aber es ist klar: Sie werden immer häufiger aufkommen.

Erst kürzlich wurde Tesla von einem US-Gericht entlastet, obwohl die KI nachweislich einen Unfall verursacht hat. Hat sich das Gericht da auch ein wenig aus der Verantwortung gezogen? Sind wir juristisch-ethisch schlicht überfordert und urteilen deshalb nach zu simplen Schemen? Wenn wir die KI nicht verstehen, dann ist eben der Mensch verantwortlich.

Nula Frei: Ja, ich sehe die Gefahr, dass da eine Art von Verantwortungs-Lücken entstehen, dass man die Schuldfrage gar nicht mehr zufriedenstellend beantworten kann. Das Problem müssen wir dringend lösen. Ich habe aber noch eine Frage zur Idee, dass KI so etwas wie einen Kolleginnenstatus haben könnte. Gäbe es denn überhaupt technische Ansätze, wie man Maschinen Ethik beibringen kann?

Denis Lalanne: Die gäbe es, ja. Aber dann müsste man wohl als erstes einmal definieren, was wir mit «ethisch» genau meinen.

Simon Ruffieux: Was zur Frage führt, ob es eine allgemeine Definition überhaupt geben kann. Oder ob die Vorstellungen von Ethik nicht wie gesagt unterschiedlich sind, je nach Kultur. Aber ja, das ist ein grosses Thema bei der Entwicklung grosser Modelle, aktuell ist wohl vor allem ChatGPT interessant: Diesem System wurde zum Beispiel beigebracht, keine Witze über Frauen zu machen.

Wie?

Simon Ruffieux: «Unmoderiert» macht diese KI natürlich Witze über Frauen, das steckt in den Daten drin. Aber in einem zweiten Schritt wurden diese Antworten geratet, von Menschen, die einem entsprechenden Ethik-Kodex

folgen, und so hat die Maschine nach und nach gelernt, solche Witze nicht mehr zu machen.

Denis Lalanne: Was eigentlich sehr spannend ist: Wir werden in Zukunft so etwas wie eine Daten-Archäologie machen können und in diesen Lernprozessen der Algorithmen ablesen können, wie es um die jeweiligen Gesellschaften stand, wie liberal, wie restriktiv, wie diskriminierend sie waren.

Dabei würde man ja bestimmt auch Kontexte unterscheiden können. Man bekommt schon ein Gefühl, wie komplex die Diskussion letztlich ist: KI im Spital und ChatGPT, das Witze macht oder sie sich verkneift – das sind ganz andere ethische Kaliber.

Denis Lalanne: Unbedingt, und es ist mir wichtig zu sagen, dass wir nicht nur über Regulation und Grenzen sprechen sollten. Manches werden wir auch nicht zu sehr regulieren wollen. Zum Beispiel hat KI in verschiedenen Feldern das Potential aufgezeigt, uns zu mehr Kreativität zu verhelfen. Dieses Potential sollten wir möglichst frei erkunden können.

Nula Frei: Im Feld der Unterhaltung braucht es wohl tatsächlich nur sehr wenig Regulierung. Wir müssen unbedingt immer den Kontext betrachten und nur da regulieren wo es echte Risiken gibt.

Simon Ruffieux: Einmal mehr, da muss ich widersprechen – so einfach ist es nicht. Denn mitunter ist es sehr schwierig vorherzusehen, wo eine KI-Technologie eingesetzt wird. Noch werden generative KI-Tools wie ChatGPT vornehmlich nur zur Unterhaltung eingesetzt, das stimmt, aber auch da schon zeigen sich negative Auswirkungen. Und wohin geht die Entwicklung noch?

Welche Gefahr sehen Sie denn, gerade im Fall generativer Modelle?

Simon Ruffieux: Ich könnte zum Beispiel ohne weiteres eine Online-Zeitung programmieren, die automatisierte Inhalte erstellt, die also von einer KI generiert ist. Das wäre gar nicht mal so aufwendig. Wenn ich das ohne Redaktion mache, wenn also niemand mehr drüberschaut, dann wird diese Zeitung ziemlich sicher Fake News sowie rassistische und frauenfeindliche Inhalte verbreiten. Und ich werde sogar noch Geld verdienen damit.

Aber von Hollywood-Szenarian à la «Terminator» sind wir ja noch weit entfernt? Sind diese Horrorszenarien einer die Menschheit ausrottenden KI überhaupt hilfreich oder verstellen sie die Sicht auf die relevanten ethischen Probleme?

Denis Lalanne: Wenn wir Risiken und Chancen gut kommunizieren, dann können wir einen Prozess anstossen. Ich mag es nicht besonders, wenn die Diskussion im Entweder-oder-Schema läuft. Ich sehe bei KI wirklich

aussergewöhnliche Chancen, wir entwickeln da gerade so etwas wie ein Super-Werkzeug. Es wäre schade, dieses Instrument zu begraben, bevor es wirklich getestet wurde, indem man versucht, die Vorteile und Nachteile zu messen. Horrorszenarien wie in Terminator? Ich sehe sie nicht wirklich – wir können unsere KIs jederzeit stoppen. Oder etwa nicht?

Simon Ruffieux: Noch können wir das, klar. Und genau deshalb ist es so wichtig, diese Kontrollmechanismen jetzt einzubauen. Denn die Risiken werden auf jeden Fall greifbarer. Zum Beispiel wenn man ganze Produktzyklen automatisiert, was bei Medikamenten gerade passiert. Wenn das System offen angelegt ist, kann man es natürlich mit der Suche nach dem effektivsten Wirkstoff betrauen. Man kann aber auch fragen, ob es die tödlichste toxische Substanz der Welt synthetisieren kann. Wenn wir da keine Wertvorstellungen in die Systeme einbauen, dann kann das gefährlich sein. Aber natürlich, die KI wird nicht selbst eine böse Absicht entwickeln, es kommt am Ende immer noch auf die Nutzerinnen an.

Nula Frei: Ich bin auch immer ein wenig skeptisch, wenn es zu sehr dramatisiert wird. Damit wir alle eine ernsthafte Diskussion führen können, ohne Panik zu verbreiten – aber auch ohne die Risiken zu unterschätzen, sehe ich nur einen Weg: Wir müssen unbedingt auf mehr «AI Literacy» in der Gesellschaft hinwirken.

Roland Fischer ist freischaffender Wissenschaftsjournalist.

Unsere Expertin ► Nula Frei ist Lehrund Forschungsrätin am Departement für internationales Recht und Handelsrecht der Universität Freiburg. nula.frei@unifr.ch



Unser Experte ➤ **Denis Lalanne** ist Professor für Informatik am Departement für Informatik und Co-Direktor des Human IST Institute. denis.lalanne@unifr.ch



Unser Experte ► Simon Ruffieux ist Lehrbeauftragter am Departement für Informatik und Senior Researcher am Human IST Institute. simon.ruffieux@unifr.ch





Du phénomène technologique à l'objet social

Au fil des années, l'intelligence artificielle a quitté les rubriques technologiques des médias pour migrer vers d'autres pages, notamment société. Désormais véritable objet social, l'IA pose de nombreux défis en matière de gouvernance et d'éthique, constate la chercheuse de l'Unifr Anna Jobin. Patricia Michaud

«Dès qu'on essaie de définir l'intelligence artificielle, il devient évident qu'il s'agit d'un objet social», selon Anna Jobin. La chercheuse et maître-assistante à l'Institut Human-IST de l'Unifr s'explique: «Actuellement, il n'y a pas de définition de l'IA qui fasse consensus; cela déclenche forcément une petite alerte dans la tête de la chercheuse en sciences sociales que je suis...» Depuis le milieu des années 1950 et le célèbre atelier «Dartmouth Summer Research Project on Artificial Intelligence» – qu'on considère généralement comme l'évènement fondateur de l'IA en tant que domaine de recherche sous ce nom -, les contours de la discipline n'ont cessé d'évoluer. «Le débat sur la question de la définition reste ouvert, que ce soit dans les milieux de l'informatique, de la politique, de l'économie ou encore des médias». Car justement, l'une des particularités de l'IA, c'est qu'elle englobe de plus en plus de pans de la société.

Mais comment est-on passé de l'intelligence artificielle en tant que phénomène essentiellement technologique à l'IA en tant que phénomène social (et politique)? C'est l'une des questions qu'aborde «Shaping AI», un projet de recherche multinational et multidisciplinaire auquel participe Anna Jobin, qui examine l'évolution du discours public sur l'intelligence artificielle dans quatre pays (Allemagne, Royaume-Uni, Canada et France) sur la période 2012–2021. «L'étude n'est pas encore terminée, mais on peut déjà dégager quelques pistes», relève la Docteure en sciences sociales qui, parallèlement à ses activités à l'Unifr, exerce la fonction de chercheuse senior auprès de l'Alexander von Humboldt Institute for Internet and Society, à Berlin.

Des discriminations perpétuées

Au nombre des éléments qui ont contribué à faire basculer l'IA dans la grande famille des thèmes de société figurent plusieurs scandales et controverses récents. «Aux Etats-Unis, un dossier journalistique publié par ProPublica, et intitulé «Machine Bias», a montré à quel point certains

programmes informatiques – basés sur l'IA – utilisés dans le système judiciaire et policier américain pour identifier de potentiels criminels peuvent être discriminatoires, voire arbitraires», relève Anna Jobin. Un autre projet américain baptisé «Gender Shades», mené par le MIT Media Lab, est pour sa part parvenu à la conclusion que certaines technologies de reconnaissance faciale sont drastiquement moins fiables dans le cas de femmes à la peau foncée que d'hommes à la peau claire. Ce n'est pas qu'une question de données de formation. «Loin de repartir à zéro, de nombreux logiciels basés sur l'IA reprennent les calibrages utilisés précédemment avec les technologies analogiques, ce qui contribue à perpétuer les discriminations.»

L'équipe de «Shaping AI» a par ailleurs constaté une évolution du traitement médiatique de l'intelligence artificielle ces dix dernières années. «Elle a de plus en plus tendance à être abordée d'un point de vue social et à migrer des rubriques technologiques vers d'autres rubriques, notamment société, politique ou encore économie.» A titre personnel aussi, Anna Jobin observe un déplacement de l'IA d'un objet purement technologique vers un objet social. «Prenons l'exemple de mes études universitaires: il y a une vingtaine d'années, j'ai dû obtenir une dérogation pour étudier simultanément la sociologie et l'informatique. Désormais, le numérique est partout; je suis invitée à présenter certaines thématiques liées à l'IA dans des cours de sociologie, de droit ou de sciences politiques.»

La gouvernance au centre

Logiquement, lorsqu'on aborde l'intelligence artificielle par le prisme de ses composantes sociales, le focus n'est pas le même que dans une approche purement technologique. «Actuellement, la question de la gouvernance est l'une des thématiques les plus saillantes en lien avec l'IA», constate la chercheuse. Après une prolifération de réglementations non contraignantes, telles que des directives et des principes éthiques, un changement s'opère. Côté législation,

l'Union européenne finalise une loi sur l'IA, qui sera la première régulation internationale du genre au monde. Priorité affichée du Parlement européen: veiller à ce que les systèmes d'IA utilisés dans l'UE soient sûrs, transparents, traçables et non-discriminatoires. Quant au Conseil de l'Europe, il s'est doté d'un comité sur l'IA. Présidé actuellement par le Suisse Thomas Schneider, il est chargé d'établir une convention cadre sur l'intelligence artificielle, les droits de l'homme, la démocratie et l'Etat de droit. «Ces règlementations auront un impact concret sur le terrain, notamment au niveau des pratiques économiques.»

«Du point de vue des sciences sociales, il est pertinent de se demander qui a le pouvoir de dicter sa vision du monde.» Est-il possible de prévenir et/ou de contrer de nouvelles inégalités, qui viendront se greffer sur les inégalités existantes? Anna Jobin rappelle notamment que «pour les petites entreprises de nombreux secteurs économiques, il devient compliqué de rester compétitives face aux grandes qui sont en mesure d'amasser d'énormes quantités de données et d'exploiter ces dernières grâce à l'IA». Selon l'observatrice, la gouvernance en matière d'intelligence artificielle est confrontée à plusieurs pièges. «Un premier risque est de se perdre dans des nuances technologiques dont on n'a pas forcément besoin pour mieux gouverner notre société.» Un autre risque, c'est que les intérêts particuliers d'acteurs trices individuel les - plutôt que les intérêts généraux de la société - guident les réflexions.

Pas si verte que ça

Au-delà des questions de régulation, l'explosion du recours à l'IA en pose de nombreuses autres, dont plusieurs de nature éthique. La question des ressources est l'une d'entre elles. «On assiste à une véritable «AI race», ou course à l'IA, entre les plus importantes nations industrialisées, telles que la Chine et les Etats-Unis; plus on est grand, plus on a accès aux *data centers*, etc.» Or, du côté des pays à la taille plus modeste, on se demande comment rester compétitif face à ces mastodontes de la technologie. «Récemment, j'étais en déplacement professionnel en Indonésie, où j'ai constaté que cette crainte était bien réelle, comme en Suisse d'ailleurs», rapporte l'observatrice. Avant d'ajouter que «s'il s'agit vraiment d'une course, personne ne la gagnera jamais puisqu'il n'y a pas de ligne d'arrivée».

Un autre aspect de l'IA qui fait couler de plus en plus d'encre est son impact environnemental. «Selon une étude dont les résultats n'ont pas encore été évalués par les pairs, il faut près d'un demi-litre d'eau pour refroidir un centre de données lors d'une conversation aller-retour avec ChatGPT, ce qui constitue une empreinte écologique assez effrayante!» Une pesée des intérêts – consistant à se demander dans quels cas l'utilisation de l'intelligence artificielle constitue vraiment une valeur ajoutée – doit donc absolument être opérée afin de limiter la casse. Il y a quelques années, une

recherche conduite par une équipe de l'EPFZ – dont faisait partie Anna Jobin – a consisté à analyser des dizaines de chartes éthiques émanant d'organisations diverses et portant sur l'IA. Alors que des principes comme la transparence et la sphère privée, ou encore la justice et la non-discrimination, étaient présents dans la plupart des documents, «nous avons été surpris de constater que d'autres principes fondamentaux en étaient quasiment absents, notamment la durabilité et l'impact sur l'environnement».

Rares aussi étaient les chartes évoquant les conséquences de l'IA sur les systèmes de protection sociale ou la solidarité. Or, «la fracture Nord-Sud est bien là dans le domaine de l'intelligence artificielle, et pas seulement au niveau des matières premières». Et la chercheuse de citer le cas de celles et ceux qu'on nomme les «travailleurs du clic», ces personnes qui passent leurs jours (et souvent leurs nuits) derrière un ordinateur, à effectuer des tâches répétitives telles qu'identifier des motifs sur des images ou à mettre en ordre des données, ce pour un salaire souvent dérisoire. Il existe donc un décalage entre la vision euphorique des nouvelles technologies en général – et de l'IA en particulier – comme parfaitement vertes et égalitaires et la réalité, plus nuancée, constate Anna Jobin.

«On en revient à l'importance d'une gouvernance conçue de sorte à ce que l'IA soit utilisée pour le bien public, sans excès ni déséquilibres», conclut la docteure en sciences sociales. Encadrer et prévenir les dommages, voilà la clé. Le hic? «Certains de ces dommages sont invisibles dans l'immédiat, car ils concernent des questions plus larges, portant sur le long terme. Quelles sont les conséquences pour la démocratie et la cohésion sociale lorsqu'il est possible de diffuser massivement – et en quelques clics – des informations erronées? Comment un système de santé est-il impacté par la possibilité d'un triage automatisé selon des critères économiques? Notre confiance dans l'Etat de droit est-elle garantie si l'on confie à l'IA la gestion de domaines tels que la jurisprudence?»

Patricia Michaud est journaliste indépendante.

Notre experte > Anna Jobin est chercheuse et maîtreassistante à l'Institut Human-IST de l'Unifr. Elle intervient aussi dans le cadre du nouveau cursus de master Digital Society, une formation interdisciplinaire unique en Suisse, qui aborde la transformation numérique sous l'angle des sciences sociales. Parallèlement, Anna Jobin exerce des activités au sein de l'Alexander von Humboldt Institute for Internet and Society, à Berlin, et le mandat de présidente de la Commission fédérale des médias. anna.jobin@unifr.ch

The Singularity is Near!

Glaubt man den Post- und Transhumanisten, so ist das Ende der Menschheit nahe: Schon in wenigen Jahren soll der Mensch von seinen eigenen, künstlich intelligenten Geschöpfen aus der Geschichte der Evolution gekickt werden. Oliver Krüger

Vor genau 30 Jahren prophezeite der amerikanische Science-Fiction-Schriftsteller Vernor Vinge bereits die Verwirklichung der so genannten Singularität: «Within thirty years, we will have the technological means to create superhuman intelligence. Shortly after, the human era will be ended.»

Ominöser Singular

Das Konzept der Singularität stammt ursprünglich aus der Mathematik und Kosmologie, wurde aber von Vinge und seinem Landsmann, dem IT-Unternehmer Ray Kurzweil, als Moment eines unumkehrbaren technologischen Durchbruches in der Entwicklung von künstlicher Intelligenz (KI) adaptiert und in Büchern und Filmen popularisiert. Mit dem von Kurzweil prognostizierten Datum im Jahr 2045 verbindet sich die Hoffnung auf die Lösung ALLER Menschheitsprobleme wie Alter, Krankheit, Tod, Armut und Überbevölkerung etc.

Bemerkenswert an dieser Zukunftsvision ist, dass stets von der einen künstlichen Intelligenz gesprochen wird, die in Kürze alles natürliche Denken transzendieren werde. Der Philosoph Yuval Harari vermutet gar, dass der Glaube an die eine, erlösende KI zur Ausbildung einer neuen Religion führen werde. Ganz offensichtlich aber entspricht der Singular der KI nicht dem tatsächlichen technischen Fortschritt, bei dem wir stets von einer Pluralität verschiedener Computerprogramme ausgehen müssen, die bestimmte Aufgaben erfüllen. An dieser Beobachtung zeigt sich, wie sehr es kulturelle Ideen sind, mit denen der recht diffuse Begriff der KI – im Singular – aufgeladen wird. Es erscheint daher vielversprechend, sich dieser Kulturgeschichte künstlicher Intelligenz und ihrem normativen Gehalt zu widmen.

Schachmatt

Der britische Mathematiker Irving John Good gilt als Urheber der Idee einer computertechnischen Superintelligenz,

die durch eine «Intelligenzexplosion» aus dem Vorhandensein von quasi unbegrenztem Speicher und extrem schnellen Prozessoren emergieren würde. Good war ab 1941 in Bletchley Park an der Entwicklung des ersten elektronischen Computers, des Colossus, unter der Leitung von Alan Turing beteiligt und wirkte später an der Virginia Tech University in den Vereinigten Staaten. Seinen berühmten Aufsatz von 1965 «Speculations Concerning the First Ultraintelligent Machine» leitet Good mit einem prophetischen Bekenntnis ein: «The survival of man depends on the early construction of an ultra-intelligent machine.» Dieser Computer, den Good bis zum Ende des 20. Jahrhunderts erwartet hatte, werde dem Menschen in der Speicherung und Verarbeitung von Informationen weit überlegen sein. Als erste Anwendung und damit als Massstab zur Messung der Intelligenz führte er das Schachspiel ein, das er selbst mit Passion betrieb. Diese Wahl war folgenreich, da nun bis zum Sieg des IBM-Computers Deep Blue über den Weltmeister Garri Kasparov im Jahr 1996 Schach als Indikator für die Verwirklichung einer dem Menschen ebenbürtigen KI galt.

Good verband mit den Supercomputern allerdings auch eine politische Vision: Schon 1962 – auf dem Höhepunkt der Kuba-Krise – spricht er davon, dass die künftigen ultraintelligenten Maschinen der Russen und Amerikaner zu einer einzigen Weltregierung verschmelzen und einen dauerhaften Frieden garantieren könnten: «Orakel aller Lander vereinigt euch!» war seine Devise.

Von Science zu Science-Fiction

Gleich in zwei bedeutenden Fällen waren Irving Goods Ideen nun richtungsweisend für populäre Werke der Science-Fiction, die bis zu den heutigen Techno-Utopien der Post- und Transhumanisten unsere Vorstellungen von einer überlegenden Computerintelligenz prägen. Da ist zum einen der alles beherrschende Bordrechner HAL 9000 eines Forschungsraumschiffes in Stanley Kubricks Film

2001: A Space Odyssey. Das filmische Meisterwerk von 1968 entstand parallel zu dem gleichnamigen Roman von Arthur C. Clarke. Dieser vermenschlichte Computer, der ohne Körper nur durch sein rotes Kameraauge das Schiff überwacht, ist schliesslich so sehr von seiner eigenen Perfektion überzeugt, dass er sich keine Fehler eingestehen kann und beginnt, die menschlichen Astronauten zu ermorden. Der Film endet damit, dass es dem letzten Überlebenden der fünf Astronauten gelingt, den Computer abzuschalten.

Die gravierendste Konsequenz dieser Inszenierung einer scheinbar weltenthobenen Superintelligenz ist die Ausblendung des menschlichen Faktors in der realen KI-Entwicklung unserer Tage

Lange bevor in jeder Vorstadtfamilie wenigstens ein Commodore-64-Homecomputer die Jugendlichen begeisterte, verkörperte der Computer HAL ein nichtmenschliches und nichtgöttliches Gegenüber, das in seiner Macht (innerhalb des Raumschiffes) den Menschen weit übertraf. Die Kybernetiker Marvin Minsky und eben Irving Good fungierten nicht nur als wissenschaftliche Berater des Regisseurs, sondern wurden bereits in der Romanvorlage von Clarke als Erfinder neuronaler Netze in der fiktiven Computerwissenschaft der 1980er Jahre gepriesen.

Auch der polnische SF-Autor Stanisław Lem bezieht sich in seiner Erzählung GOLEM XIV (dt.: Also sprach Golem) von 1981 auf Good. Eigentlich für militärische Zwecke ersonnen, wurde der Supercomputer GOLEM XIV nach einigen misslungenen Tests zum Massachusetts Institute of Technology in Boston abgestellt, wo er mehrere philosophische Vorlesungen über den Menschen, sich selbst und den Sinn des Lebens hält (die Vorlesungen werden fiktiv von «Irving T. Creve» ediert). In seiner ersten anthropologischen These konstatiert das Elektronenhirn, dass der Sinn des Menschen nur seine «Botschaft» sei, da die Organismen deren Übermittlung dienen würden und nicht die «Botschaft» den Organismen. Aus Sicht des GO-LEM stellt der Mensch daher ein Übergangswesen dar, das seinen Sinn lediglich darin habe, wirklich «vernunftbegabte Wesen», nämlich Computer, zu bauen.

Unschwer lassen sich in diesen visionären und populären Bildern der künstlichen Intelligenz die Konturen des europäischen Geniebegriffes des späten 19. Jahrhunderts erkennen. Unabhängig von aller Parteilichkeit würde die schiere Masse an vorhandener Intelligenz alle Weltprobleme durch die überwältigende Kraft einer rationalen Vernunft lösen können – es ist deshalb kein Zufall, dass Schach so häufig als Indikator für ebendiese geniehafte Intelligenz herangezogen wurde.

Die gravierendste Konsequenz dieser Inszenierung einer scheinbar weltenthobenen Superintelligenz ist die Ausblendung des menschlichen Faktors in der realen KI-Entwicklung unserer Tage. Als es zu Beginn der 2000er kaum Fortschritte in der Forschung gab, verhalf eine immense Erweiterung der visuellen und textlichen Datenmengen, diese Schwierigkeiten zu beheben. Diese millionenfachen Daten müssen jedoch im Prozess des data labeling zunächst von Menschen zugeordnet werden, z.B. dass auf einem Bild eine tierische Maus zu erkennen ist, selbst wenn sie verletzt oder tot ist. Auch Kommunikationsprogramme wie Chat-GPT (Generative Pre-trained Transformer) beruhen auf diesem umfangreichen Training mit Hilfe menschlicher Akteure, damit die KI z.B. keine anstössigen Gewalt- oder Sexfantasien produziert. Zwar werden auch die verbreiteten Captcha-Aufgaben zum Training von KI genutzt, in der Regel aber werden hier data worker vor allem aus dem globalen Süden zu Dumpinglöhnen eingesetzt. Eine psychologische Betreuung für dieses digitale Sortieren von erwünschten wie unerwünschten menschlichen Verhaltensweisen im Auftrag globaler IT-Konzerne wird kaum geleistet.

Alles andere als autonom

Während im technophilen Diskurs des Westens über die Singularität und die Chancen und Risiken der KI für die sehr abstrakte «Zukunft der Menschheit» debattiert wird, bringt das ganz reale KI-Training ein weltweites, digitales Prekariat hervor. Die Mystifizierung einer autonom erscheinenden KI erweist sich dabei als äusserst hilfreich, um die ausbeuterischen ökonomischen Strukturen zu verschleiern. Eine ethische Debatte über KI muss an diesem Punkt ansetzen.

Unser Experte ► Oliver Krüger ist Professor für Religionswissenschaft im Departement für Sozialwissenschaften. oliver.krueger@unifr.ch

Literatur

- ▶ Phoebe Moore & Jamie Woodcock: Augmented Exploitation. Artificial Intelligence, Automation and Work. London: Pluto Press 2023.
- ▶ Adrienne Williams & Milagros Miceli & Timnit Gebru: Artificial Intelligence. In: Noēma 13.10.2022.
- ▶ Oliver Krüger: Virtual Immortality. God, Evolution, and the Singularity in Post- and Transhumanism. Bielefeld: transcript 2021.





Label für digitale Nachhaltigkeit

In disziplinenübergreifenden Arbeiten wollen Fachkundige aus der Umweltphilosophie und aus der Informatik die ökologischen wie auch die sozialen Auswirkungen von digitalen Dienstleistungen abschätzen – und sie mit ethisch vertretbaren Grenzwerten in Beziehung setzen. Ori Schipper

Wenn ihr Vorhaben eine Partitur wäre, kämen die Paukenschlägerinnen und -schläger voll zum Zug. In ihren Fachbeiträgen geht es um nichts weniger als um die «Bewältigung von ethischen Herausforderungen im Zusammenhang mit künstlicher Intelligenz». Und helfen soll dabei das «Freiburger Modell für eine nachhaltige digitale Ethik», das ein Team um Ivo Wallimann-Helmer, Professor für Umweltgeisteswissenschaften am Departement für Geowissenschaften, und Edy Portmann, Professor am Departement für Informatik, seit etwa zwei Jahren am Erarbeiten ist.

Im gemeinsamen Gespräch gehen die beiden darauf ein, was dieses Freiburger Modell ausmacht – und worauf es abzielt. «Wir wollen mit unserem inter- und transdisziplinären Vorhaben Nachhaltigkeitsstandards für Algorithmen definieren», sagt Wallimann-Helmer. «Und dann in einem zweiten Schritt diese Standards zusammen mit Partnern aus Industrie und Verwaltung in der Praxis testen und etablieren», fährt Portmann fort.

Dass die zunehmende Digitalisierung dazu führt, dass sich auch die ökologischen Auswirkungen der verschiedenen Anwendungen von künstlicher Intelligenz (KI) intensivieren, liegt eigentlich auf der Hand. Doch weil eine bestimmte Anwendung sowohl positive wie auch negative Folgen für die Umwelt nach sich ziehen kann, ist deren Bewertung alles andere als trivial.

Als Beispiel führt Portmann ein Programm an, das die Paketverteildienste darin unterstützt, optimale Routen für die Auslieferung zu finden. Dank solchen optimierten Routen müssen die Mitarbeitenden des Verteildiensts weniger Kilometer zurücklegen. Dadurch setzen ihre Fahrzeuge weniger gesundheitsschädlichen Feinstaub und weniger Kohlendioxid frei – oder brauchen weniger elektrische Energie, wenn das Motorrad oder der Kleinbus von Batterien angetrieben wird. Gleichzeitig aber gilt es zu berücksichtigen, dass die für die Berechnung der optimalen Routen eingesetzten Rechner auch Energie benötigen.

Und dass die Hersteller der Rechner etwa auch auf seltene Erden angewiesen sind, die nur mit grossem Aufwand aus Minen gewonnen werden, die die Umwelt belasten.

«Wir möchten aufzeigen, in welchen Bereichen eine Entwicklung möglich ist, die weder die von der Ökologie vorgegebenen planetarischen Grenzen überschreitet, noch die sozialen Grenzen der Gerechtigkeit untergräbt»

Ivo Wallimann-Helmer

Für ihr «Freiburger Modell für eine nachhaltige digitale Ethik» wollen Wallimann-Helmer und Portmann aber auch die sozialen Auswirkungen der Digitalisierung berücksichtigen. Deshalb lehnen sie sich an die erstmals vor zehn Jahren veröffentlichten Ideen an, die der so genannten Donut-Ökonomie zugrunde liegen: Sie gehen von einem Freiraum für die Menschheit aus, der – wie das runde Süssgebäck mit dem Loch in der Mitte – sowohl nach innen wie auch nach aussen begrenzt ist.

Während der äussere Rand für die verschiedenen Belastungsgrenzen der Erde steht (wie etwa das Artensterben, die Versauerung der Ozeane oder die Abholzung der weltweiten Wälder), symbolisiert der innere Rand die gesellschaftlichen Folgen der Digitalisierung (also etwa die Transformation der Arbeitswelt, die Neuaufteilung des Einkommens, oder eine drohende Ungleichbehandlung, Marginalisierung oder Diskriminierung bestimmter Bevölkerungsgruppen). «Wir möchten aufzeigen, in welchen Bereichen eine Entwicklung möglich ist, die weder die von der Ökologie vorgegebenen planetarischen Grenzen überschreitet, noch die sozialen Grenzen der Gerechtigkeit untergräbt», sagt Wallimann-Helmer.

Digitalisierungsprojekte, die in diesem nachhaltigen Freiraum stattfinden, könnten mit einem Label gekennzeichnet werden, das die Forschenden um Wallimann-Helmer und Portmann im Austausch mit Politik und Wirtschaft schaffen wollen. Wie andere Nachhaltigkeitszertifikate (etwa für das Essensangebot von Kantinen oder für die Bewertung von Gebäuden) könnte auch das Label für nachhaltige Digitalisierung in verschiedene Stufen unterteilt sein. «Wir könnten unterschiedlich nachhaltigen Projekten zum Beispiel Gold-, Silber- und Bronzemedaillen verleihen», führt Wallimann-Helmer aus.

Dass die künstliche Intelligenz uns Menschen schon bald überflüssig macht, befürchtet keiner der beiden Gesprächspartner. «Das Herzstück der heutigen KI sind immer noch die Daten, mit der sie gefüttert wird», sagt Portmann. «Und diese Daten werden von Menschen erstellt, deshalb ist die künstliche Intelligenz auf uns angewiesen», fügt Wallimann-Helmer hinzu.

«Die Herausforderung liegt nicht im Verfassen solcher Richtlinien, sondern in deren Umsetzung» Edi Portmann

Doch wenn die Datensätze falsche, veraltete, unvollständige oder auch verzerrte Daten enthalten, können die damit gefütterten Algorithmen «die in unserer Gesellschaft bereits bestehenden Ungleichheiten etwa in Bezug auf Klasse, Rasse und Geschlecht aufrechterhalten und sogar verstärken», schreibt Portmann. Als Beispiel führt er ein KI-Tool an, das Amazon entwickelt hatte, um Stellenbewerbungen zu prüfen. Als sich bald darauf herausstellte, dass das Tool Frauen diskriminierte, stampfte Amazon es wieder ein.

Wie stellen sich Portmann und Wallimann-Helmer eine sozial gerechte künstliche Intelligenz vor? Auch hier greifen die beiden auf Ideen zurück, die vor wenigen Jahren veröffentlicht wurden. So verweisen sie etwa auf die «Ethik-Leitlinien für eine vertrauenswürdige künstliche Intelligenz» der Europäischen Union. Und auf die zahlreichen weiteren Richtlinien, die von Unternehmen wie Microsoft oder IBM, aber auch von Non-Profit-Organisationen wie Amnesty International erstellt worden sind. «Wenn man die vielen Dokumente miteinander vergleicht, fällt auf, dass meist die gleichen ethischen Werte wie etwa der Schutz der Privatsphäre und die Transparenz über das Vorgehen und die Ziele der Algorithmen im Vordergrund stehen», sagt Portmann. «Doch die Herausforderung liegt nicht im Verfassen solcher Richtlinien, sondern in deren Umsetzung.»

Ethische Entscheide berechenbar machen

Um die Herausforderung der Umsetzung zu meistern, möchte Portmann ethische Entscheide berechenbar machen: «Das Ziel ist, geisteswissenschaftliche Konzepte der traditionellen Ethik in konstruktive Algorithmen zu überführen», hält der Informatiker in einem letztjährig erschienenen Fachbeitrag fest. Diese «konstruktiven Algorithmen» sollen «Konversationsschlaufen durchlaufen» und mit Wörtern statt mit Zahlen rechnen können. Das wird es «einer künstlichen Intelligenz erlauben, passende moralische Entscheidungen situativ aus den vorhandenen Informationen und Rahmenbedingungen abzuleiten».

Ob sich dieses Vorhaben so durchführen lässt (und die gewünschten Resultate liefert), muss sich noch weisen. Schon heute jedoch steht fest, dass sich aus dem gezielten Einsatz von prädiktiven Technologien und der immer umfassenderen Sammlung von persönlichen Daten ein grosses Machtmissbrauchspotenzial ergibt. «In den letzten Jahrzehnten sind die Internetgiganten rasant – und weitestgehend ohne demokratische Kontrolle – gewachsen», gibt Wallimann-Helmer zu bedenken. Wie vielen weiteren Forschenden geht es den beiden auch darum, Demokratie und Rechtsstaatlichkeit im Zeitalter der künstlichen Intelligenz zu verteidigen. Oder in den Worten von Wallimann-Helmer: «Wir möchten mit unseren Arbeiten zur Diskussion beitragen, wie eine technologische Unterdrückung der Menschheit vermieden werden kann.»

Ori Schipper ist freischaffender Wissenschaftsjournalist.

Unser Experte ► Ivo Wallimann-Helmer ist Professor für Umweltgeisteswissenschaften an der Universität Freiburg.

ivo.wallimann-helmer@unifr.ch

Unser Experte ► Edy Portmann ist Leiter der Forschungsgruppe für Soft and Cognitive Computing und Professor am Departement für Informatik der Universität Freiburg. edy.portmann@unifr.ch

L'IA à l'école ou l'école de l'IA

L'intelligence artificielle a déjà forcé la porte de nos établissements de formation. Inutile d'envisager de faire sans. Cependant, il devient impératif de penser sa place dans l'éducation et la formation. Lionel Alvarez

Aujourd'hui, plus que jamais, les technologies numériques nouvelles invitent chaque partenaire de l'éducation et de la formation à se questionner. Au niveau technique, on pense évidemment au chatbot construit à l'aide de LLM (large language model), à l'instar de ChatGPT, permettant à chacun-e de converser avec une machine, par exemple, et de se laisser surprendre. En matière d'éducation et de formation, après le débat sur l'évaluation annoncée comme caduque par certaines personnes, émerge une question bien plus fondamentale: quelle relation souhaitons-nous favoriser entre l'IA et l'humain dans les contextes de formation et d'éducation?

Certes, une réponse semble admise: former à des usages efficients et à un esprit critique. Oui, évidemment. Ne répondrions-nous pas ceci pour chaque technologie nouvelle? Toutefois, ce projet éducatif donne le vertige au regard de l'évolution fulgurante des techniques numériques et de la multiplicité des outils qui se partagent (je n'ose dire qui se démocratisent, lorsqu'on regarde qui sont les propriétaires des solutions aujourd'hui populaires). Mais doit-on contraindre les curriculums de formation et d'éducation à l'émergence incessante d'outils techniques nouveaux?

De l'exploration de la cognition partagée...

Face à la limite de ma cognition humaine, je m'accompagne alors d'une cognition synthétique pour penser cette relation complexe entre apprentissage humain et augmentation technique de l'humain. Je sors un peu du cadre en n'employant délibérément pas le LLM à la mode, car il n'a d'une part pas besoin de publicité et, d'autre part, parce qu'il me semble aujourd'hui nécessaire de participer à la démocratisation d'autres solutions que celles qui sont propriétaires. Je converse alors avec *Wizard v1.1*, un LLM basé sur LLa-MA et entraîné à l'aide d'une nouvelle méthode, appelée Evol-Instruct, sur des données d'instructions complexes.

Lionel: Bonjour Wizard v1.1. Considère que tu es expert en intelligence artificielle, en éducation et formation. Tu donnes des réponses brèves, mais actuelles et liées à des sources explicitées.

Wizard v1.: Salut Bonjour! Je suis le Wizard v1.1, votre assistant IA spécialisé dans l'intelligence artificielle, l'éducation et la formation. Veuillez me dire comment je peux vous aider aujourd'hui?

Lionel: Quels sont les enjeux essentiels pour l'éducation et la formation, aujourd'hui que les IA génératives sont disponibles?

Wizard v1.1: Les enjeux essentiels pour l'éducation et la formation actuelles sont nombreux. Avec l'avènement des IA génératives, il est crucial de s'adapter aux nouvelles tendances et d'envisager les changements qui peuvent être apportés dans le domaine de l'enseignement et de la formation. Voici quelques points clés à prendre en compte:

- Adaptation des programmes éducatifs: Les IA génératives offrent une nouvelle façon d'apprendre et de créer du contenu. Il est important d'adapter les programmes scolaires pour inclure l'utilisation de ces technologies tout en maintenant une approche traditionnelle à la fois pratique et innovante.
- 2. Développement des compétences numériques: Les IA génératives renforcent l'importance d'apprendre les compétences nécessaires pour interagir avec ces technologies, telles que le programmeur de langage de balisage (HTML), le langage de marquage up (XML) et le langage de balisage du web (CSS).
- 3. Enseignement des valeurs fondamentales: Les IA génératives peuvent contribuer à la diffusion d'idées erronées ou discriminatoires si elles ne sont pas correctement encadrées. Il est essentiel de mettre en place



- un enseignement qui promeut les valeurs fondamentales, telles que l'égalité, la tolérance et le respect.
- 4. Favoriser la créativité: Les IA génératives peuvent être utilisées pour stimuler la créativité des étudiants en leur fournissant un outil de production de contenu innovant. Il est important d'encourager l'utilisation de ces technologies dans le cadre d'activités éducatives et de formation professionnelle.
- 5. Mise à jour du curriculum: Les IA génératives sont une révolution majeure dans les domaines de l'intelligence artificielle, de la robotique et des sciences cognitives. Il est crucial d'ajuster le curriculum scolaire pour inclure ces nouvelles connaissances scientifiques et techniques.
- 6. Formation continue pour les enseignants: Les IA génératives nécessitent une approche pédagogique différente, mettant l'accent sur la collaboration entre humains et machines. Il est essentiel que les enseignants soient formés à ces nouvelles technologies afin de pouvoir les intégrer efficacement dans leur travail
- 7. Promouvoir une approche interdisciplinaire: Les IA génératives touchent un large éventail de domaines, tels que la linguistique, l'informatique et la psychologie. Il est important d'adopter une approche interdisciplinaire pour aborder les enjeux liés aux IA dans le contexte de l'éducation et de la formation.

En tenant compte de ces enjeux essentiels, nous pouvons travailler ensemble à développer des programmes éducatifs innovants qui préparent nos élèves et étudiants au monde changement par les IA génératives. (sic)

Ce bref échange montre ce qu'il est probable de répondre aujourd'hui. En effet, les textes générés par ces IA émergent de calculs statistiques fondés notamment sur la probabilité d'apparition d'un mot. Analysons le vocabulaire choisi et prenons le temps d'étudier les recommandations proposées, ainsi que les non-dits.

... à la posture critique

Le vocabulaire tourne régulièrement autour de tendances, de nouveautés, de révolution ou de mise à jour. La tradition est certes évoquée, mais associée à l'innovation. Il y a donc dans la réponse une forme de valorisation d'un renouveau en éducation et formation. Les mots invitent aussi à l'adaptation de l'humain à la machine: «adapter les programmes scolaires», «apprendre les compétences nécessaires», «encourager l'utilisation de ces technologies» ou encore «mettre l'accent sur la collaboration entre humains et machines». C'est présenté comme si la technique était imposée, autodéterminée et que nous, simples humains, n'avions d'autre choix que d'en faire usage. Enfin, il y a quelques éléments

d'ouverture, comme une invitation à replacer les valeurs fondamentales ou une nécessité de travailler de manière interdisciplinaire, mais cela reste marginal, trop à mon avis. Il devient alors tout à fait intrigant de se remémorer que les propositions de *Wizard v1.1* ne sont «que» le fruit des textes qui avaient été écrits pas des humains auparavant. Il semble que le rapport «technophile» était prioritaire et qu'une vision autodéterminée des techniques dominait dans les *data* ayant nourri cette IA. Où sont passés les philosophes de la technique, par exemple, ou les personnes travaillant dans le champ du *social shaping of technology*?

Quant aux non-dits... C'est là que la cognition humaine - la sagesse? - devient essentielle. Quid des avis divergents? Des propos peu référencés? Des affirmations qui, malgré un intérêt potentiel, ne sont pas apparues dans le champ des possibles «hallucinés» par Wizard v1.1? Chaque personne, avec sa culture et ses savoirs, saura proposer une analyse nouvelle que l'IA n'aura pas envisagée. Je pense premièrement au capabilisme sous-jacent au discours (est-ce mon passif de pédagogie spécialisée?). Nous, humain·e·s, devrions-nous nous adapter et nous augmenter grâce à la technique pour ne pas devenir en «situation de handicap technologique»? Ce sont les individus capables qui semblent dignes dans ce discours. Deuxièmement, la tradition et des décennies de savoirs pédagogiques semblent être dévalorisées. L'innovation prime. Troisièmement, c'est une logique de productivité qui est à l'œuvre. Or, dans l'enseignement et l'apprentissage, il est parfois pertinent de s'arrêter et de prendre le temps de penser, d'embrasser le temps long, la complexité et l'effort. Il y a tant de postures critiques à développer, selon les contextes et les usages, pour que ces IA génératives puissent trouver la place qu'elles méritent, ni plus ni moins, dans les processus complexes d'apprentissage.

Ceci nécessite d'avoir le goût de dépasser la réponse crédible ou le souhait de ne pas se contenter du probable. La différence entre apprendre et produire devra être rappelée aux logiques d'employabilité. Enfin, il y aura prochainement un possible habitus à stimuler chez chaque apprenant-e: penser par soi-même. Cette posture critique n'est pas un rejet des IA génératives, mais bien une tentative de contrebalancer des propos simplement émerveillés ou des affirmations essentiellement anxieuses pour participer à la construction d'une relation sereine entre l'éducation, la formation et les outils numériques qui émergent.

Notre expert **Lionel Alvarez** est docteur en pédagogie spécialisée, lecteur au Centre d'enseignement et de recherche pour la formation à l'enseignement au secondaire et professeur associé à la Haute école pédagogique de Fribourg. lionel.alvarez@unifr.ch

Un contrat? Adressez-vous à mon assistant IA!

En droit suisse, un contrat est en principe le fruit d'un accord de manifestation de deux (ou plusieurs) volontés, réciproques et concordantes. En d'autres termes, pour conclure un contrat, il faut être en mesure d'exprimer une «volonté» de manière expresse ou tacite. Qu'en est-il toutefois, lorsqu'intervient un assistant numérique? Pascal Pichonnaz

Depuis des années déjà, les ordres en bourse sont générés par des programmes d'ordinateurs, souvent complexes, qui reposent principalement sur une approche appelée «systèmes experts». Ces algorithmes permettent de mettre en œuvre des règles prédéfinies sur des faits pour produire des conséquences. Par exemple, si tels ou tels faits sont donnés, alors il faut acheter telles ou telles actions dans telles ou telles quantités. Dans cette hypothèse, une volonté humaine a bien fixé les règles à appliquer et, le plus souvent, délimité les données (faits) à considérer. Dès lors, la conclusion d'un achat d'actions de cette manière est le fruit d'une volonté, concrétisée au travers d'un programme algorithmique «expert». Pour le droit des contrats, il n'y a rien de très particulier.

Les enjeux du machine learning

La question juridique devient toutefois beaucoup plus délicate lorsque les faits ou données à prendre en compte ne sont plus délimités à l'avance ou que l'algorithme utilisé n'est plus un système expert, avec des règles prédéfinies, mais un système de machine learning. En effet, un tel algorithme repose sur une approche dynamique, en ce sens que les réseaux de neurones artificiels permettent d'identifier, de manière autonome, des tendances récurrentes (patterns) sur la base de l'analyse d'une très large quantité de données. Durant une phase d'apprentissage (si celle-ci est prévue ou nécessaire), les patterns reconnus sont validés ou invalidés par un être humain. Cela permet d'adapter le réseau neuronal artificiel afin de le rendre plus performant; cela permet surtout de donner le sens à attribuer aux patterns spécifiques («Oui, cette image représente bien un chien»). Certains systèmes de machine learning

plus récents sont même auto-apprenants; par la répétition et surtout la rétroanalyse, ils peuvent ainsi auto-corriger leurs résultats en modifiant les poids des divers facteurs de manière autonome, ce qui améliore le résultat global. Dans ce cas également, il faut toutefois que des règles extérieures donnent du sens aux divers *patterns*. C'était l'un des principes d'AlphaGo Zero à qui l'on avait donné, en 2017, les règles du jeu de Go. Le programme avait alors joué contre lui-même de manière autonome plus de 4,9 millions de fois en trois jours, puis avait battu 100 parties à 0 la première version d'AlphaGo Lee, version entraînée par les humains. Et pourtant, AlphaGo Lee avait battu, en 2016, Lee Sedol, champion et grand maître sud-coréen du jeu de Go.

Quand le smart fridge décide

Imaginons maintenant une machine à café qui commande automatiquement la quantité et le type de capsules de café en fonction de l'usage de son utilisateur-trice, de ses habitudes de consommation et de nombreux autres facteurs ou un réfrigérateur intelligent qui gère lui-même les commandes de nourriture. En principe, l'utilisateur trice doit entraîner le réfrigérateur intelligent pour lui indiquer, sur une période de quelques semaines ou mois, ce qu'elle ou il y entrepose, les produits utilisés, ainsi que d'autres indications déterminantes, comme la présence de personnes pour des repas spécifiques ou des fêtes, des vacances, voire la fréquence de son activité physique. Après cette phase d'apprentissage, durant laquelle le réfrigérateur intelligent procède à des commandes sous la supervision de celle ou celui qui l'utilise, il va pouvoir agir de manière autonome, anticipant les besoins sur la base des probabilités établies lors de la phase d'entraînement. L'IA a identifié des patterns et les a corrélés, tandis que la période d'apprentissage a permis de leur donner du sens – par ex. faire une commande de champagne, parce qu'il y a un anniversaire à fêter. Dès la fin de la période d'apprentissage, l'IA conclut des contrats d'achat et libère les paiements de manière autonome. Comment reconnaître alors une valeur juridique contraignante à ces contrats, dès lors qu'il n'y a pas de volonté humaine à la base de ces décisions? Que se passe-t-il si les commandes sont erronées par rapport aux besoins?

C'est pour répondre notamment à ces questions que le European Law Institute, un *think tank* indépendant, soutenu notamment par la Commission européenne, a décidé de consacrer l'un de ses projets aux contrats algorithmiques. L'idée est de poser des principes et des règles-modèles pour assurer une certaine prévisibilité dans le traitement des principales questions juridiques qui se posent.

Dans l'exemple du réfrigérateur intelligent, une question juridique débattue est celle de savoir comment attribuer une «décision algorithmique» (appelé en anglais, automated decision making, ADM) à une personne. L'intelligence artificielle utilisée, que l'on peut appeler assistant numérique, est-elle un auxiliaire, un représentant, un substitut ou un simple outil prolongeant l'intention et la volonté de son utilisateur·trice? Ce n'est pas le lieu d'analyser cette question, sauf à souligner que les divers régimes juridiques en Europe devront admettre, d'une manière ou d'une autre, que tout contrat passé avec un assistant numérique est un contrat passé par son utilisateur trice. En effet, en acquérant un réfrigérateur intelligent et en l'entraînant, puis en lui donnant l'autonomie requise, on consent à l'interaction de cet assistant numérique avec des fournisseurs (euxmêmes d'ailleurs peut-être assistés d'assistants numériques) et au risque qu'elle ou il prend en l'utilisant.

Le consentement et le droit d'objecter

En droit, pour produire un effet, le consentement doit être éclairé. Par exemple, en matière médicale, cela signifie que la patiente ou le patient doit être informé-e sur les risques d'une opération, ses conséquences financières et, en particulier, sur sa couverture par une assurance. Pour notre assistant numérique, il faut dès lors se demander si, lors de l'achat du réfrigérateur intelligent, celle ou celui qui achète a été suffisamment informé-e sur son fonctionnement, le type d'algorithme et les risques liés à un assistant numérique. Il faut donc que la législation ou du moins les tribunaux posent un certain nombre d'exigences à ce devoir d'information au moment de la vente.

Cela dit, bien que dûment informé, l'utilisateur trice voudra peut-être pouvoir intervenir pour approuver ou objecter à un contrat conclu par l'assistant numérique. En effet, considérant que le nombre de calories par jour est supérieur à ce qui est souhaitable pour une alimentation saine, compte tenu de la probabilité du nombre de personnes

se servant des produits du réfrigérateur et des données d'activité fournies par la montre connectée de celui ou celle-ci, l'assistant numérique pourrait n'acquérir plus que du beurre allégé ou modifier de manière significative le régime alimentaire tout en respectant les goûts du ou de la titulaire. Faut-il dès lors prévoir un droit à objecter aux contrats conclus par l'assistant intelligent ou l'obligation

En droit, pour produire un effet, le consentement doit être éclairé

d'approuver tout contrat? La seconde option n'est pas réaliste, puisque le but même de l'acquisition d'un assistant numérique est d'automatiser les contrats. Néanmoins, on pourrait envisager un droit d'objecter; ce qui supposerait que l'assistant numérique informe l'utilisateur-trice en cas de changement significatif du type d'achats. Il faudrait alors évidemment déterminer les paramètres du «changement significatif», mais peut-être aussi la possibilité de pouvoir les revoir en cours d'utilisation.

Des questions contractuelles à revoir

Enfin, qu'en est-il des conditions générales et des clauses contractuelles imposées par le fournisseur? Doivent-elles être lisibles sous forme numérique pour être prises en compte? Peut-on prévoir un alignement des conditions générales sur des standards prédéterminés, sous peine d'exclure les contrats avec tel ou tel fournisseur? Peut-on appliquer un régime de clauses abusives pour contrats numériques? En d'autres termes, comment s'assurer que la protection des consommateurs-trices, développée dans un monde analogique, puisse continuer à s'appliquer dans un monde où les assistants numériques seront très présents?

Le European Law Institute tente notamment d'établir si l'acquis communautaire en matière de protection des consommateurs trices est prêt pour les contrats algorithmiques et, bien sûr, de déterminer les modifications à apporter pour répondre aux besoins légitimes des consommateurs trices en Europe, et – espérons-le – aussi en Suisse.

Notre expert > Le Professeur Pascal Pichonnaz est titulaire de la Chaire de droit privé et de droit romain, ainsi que président du European Law Institute (ELI), www.europeanlawinstitute.eu pascal.pichonnaz@unifr.ch



Asymmetrie zwischen Lob und Tadel

Wer ist schuld, wenn ein autonomes Fahrzeug eine Fussgängerin nicht erkennt oder wenn ein Chatbot rassistische Äusserungen tätigt? Ganz klar: der Mensch. Aber wie steht es, wenn eine KI etwas Tolles hervorbringt? sarah Köglsperger

Auch wenn die gewisse Eigenständigkeit von KI-Systemen die Verantwortungszuschreibung verkompliziert, lautet die gängige Ansicht, dass dennoch menschliche Akteure – meist nachlässige Nutzer_innen – zur Verantwortung gezogen werden können, wenn eine künstliche Intelligenz versagt oder Fehler macht.

Erst seit Kurzem wird auch über die positive Seite der Verantwortung diskutiert: Wer verdient Lob und Anerkennung für Leistungen, die mithilfe von KI erzielt wurden? Generell scheint eine Asymmetrie zwischen Lob und Tadel in unserer moralischen Praxis vorzuliegen. Wir neigen dazu, Menschen schnell für schlechtes Verhalten verantwortlich zu machen, selbst bei kleineren moralischen Fehltritten oder Nachlässigkeiten. Hingegen reicht eine geringfügig positive Handlung oft nicht aus, um Anerkennung zu erhalten. Stattdessen braucht es besondere Leistung, Talent oder Anstrengung. Angewandt auf den Kontext von KI, bedeutet das, dass wir zwar menschliche Nutzer_innen für die Fehler von KI-Systemen zur Verantwortung ziehen, diesen jedoch keine oder nur geringfügige Anerkennung zukommen lassen, wenn jemand etwa mithilfe von ChatGPT ein hervorragendes Essay geschrieben hat.

Im universitären Kontext, in dem schriftliche Arbeiten nach wie vor eine zentrale Rolle in der Leistungsbeurteilung spielen, stellt uns das vor die Frage, wie wir mit mithilfe von Large Language Models – kurz LLMs – verfassten Arbeiten umgehen sollten. Eine mögliche Antwort wäre, Studierende weiterhin als verantwortlich zu betrachten, da sie aufgrund der Limitationen von LLMs immer noch eine hinreichend bedeutsame Aufgabe im Schreibprozess übernehmen. LLMs werden auf riesigen Textkorpora trainiert,

mit dem Ziel, dass sie Muster in der Sprache erkennen und als Antwort auf einen Prompt des Nutzers eine passende Textausgabe produzieren können, indem sie die wahrscheinlichsten Wörter in einem Kontext vorhersagen. Das bedeutet, dass LLMs vielleicht praktikabel sind, um auf Emails zu antworten oder um der Grossmutter eine Dankeskarte zu verfassen, aber sie beim wissenschaftlichen Schreiben an gewisse Grenzen stossen. Denn, da sie Texte auf Basis der Häufigkeit von Wörtern in einem Kontext generieren, besteht die Gefahr, dass sie zwar glaubwürdig erscheinende, jedoch inkorrekte Informationen liefern. (Wer schon mal ChatGPT nach Literaturvorschlägen oder Quellenangaben gefragt hat, kennt die Problematik.) Spezifisches Fachwissen oder neuere Geschehnisse, die nicht in den Trainingsdaten enthalten sind, stellen ebenfalls Limitationen dar, genauso wie die Entwicklung neuartiger Argumente und Positionen. Nutzer_innen müssen daher immer noch die ausgegebenen Informationen überprüfen, fachspezifische Debatten selbst recherchieren, und eigene Position entwickeln.

Ist die Lob-Tadel Asymmetrie beim Einsatz von LLMs also gar nicht relevant, da der Nutzer immer noch die zentrale Arbeit leistet? Man kann hier zurecht einwenden, dass sich diese Darstellung nur auf die Funktionen jetziger Sprachmodelle bezieht. Ein Grossteil der erwähnten Limitationen wird hingegen bereits mit der Veröffentlichung von GPT-4 wegfallen. Dieses wird Plugins haben, das unter anderem auch Internetsuchen erlaubt, wodurch der Chatbot auf aktuelle Informationen zugreifen kann, ohne dass sie in den Trainingsdaten enthalten sein müssen.

Ausserdem kann es zwar sein, dass ChatGPT in Zukunft noch nicht die Kernaufgaben beim wissenschaftlichen Schreiben übernehmen wird, jedoch sagt uns das nichts über die Verantwortung für studentische Arbeiten. Vor allem zu Studienbeginn haben schriftliche Arbeiten oft das Ziel, zu demonstrieren, dass die Seminarinhalte und die wichtigsten Positionen in der Literatur verstanden wurden. Und hierfür kann ChatGPT hervorragend geeignet sein. Mit einem präzise formulierten Prompt kann es eine passende Fragestellung mit möglicher Gliederung vorschlagen sowie einen ersten Textentwurf generieren. Nutzer_innen haben nur noch die Aufgabe, passende Prompts zu formulieren und den generierten Text zu überprüfen.

Genügt das, um Anerkennung für eine gute Seminararbeit zu verdienen? Hierfür sollten wir uns klar machen, was es an einer guten Arbeit ist, das wir durch Lob, respektive einer guten Note, honorieren. Zum einen bezieht sich das Lob auf die Schreibfähigkeiten im engeren Sinn, deren Erwerb viel Zeit und Übung erfordert und die es erlauben, den eigenen Gedankengang für den Leser verständlich und leicht nachvollziehbar zu präsentieren. Zum anderen bezieht sich das Lob auf die tatsächlichen Aussagen des Autors. Diese sollten eigenständiges Denken sowie Kreativität und die Fähigkeit, selbst Stellung zu beziehen, zeigen. Zur Illustration bietet sich der Vergleich zu einem Kunstwerk an: Was wir an einem grossartigen Gemälde loben, sind einerseits die künstlerischen Fähigkeiten, die für die Schaffung nötig waren – im Sinn vom griechischen (technê) - und andererseits die kommunikative Dimension des Kunstwerks. Der Künstler konstruiert das Gemälde nach seinen Vorstellungen, um eine Idee oder ein Gefühl zu vermitteln oder die Betrachterin zum Nachdenken anzuregen. Wie der Künstler jeden Pinselstrich bewusst setzt, wählt auch die Autorin oder der Autor jedes Wort bewusst aus, um seine Gedanken der Leserschaft möglichst akkurat zu vermitteln. Aber es ist nicht nur so, dass wir das Schreiben rein als kommunikatives Mittel nutzen, sondern das Schreiben hilft uns. unsere Gedanken zu ordnen. Schreiben ist Denken - was der Begriff (logos) passend bezeichnet, der nicht nur (Wort) und (Sprache), sondern auch (Überlegung) und «Vernunft» heissen kann.

Gewusst wie

Wir benötigen also mit ChatGPT nicht nur geringere Schreibfertigkeiten, sondern auch intellektuelle Fähigkeiten wie kritisches, eigenständiges Denken werden nicht mehr (in dem Umfang) kultiviert, wenn wir Text generieren lassen. Wir sollten uns auch klar sein, dass, da LLMs Sätze basierend auf der Häufigkeit von Wörtern generieren, es keine Intentionalität und keinen Gedanken gibt, der durch den Text vermittelt werden soll. Wenn bei schriftlichen Arbeiten zuvor der Gedankengang des Autors und dessen präzise Darstellung evaluiert wurde, wird jetzt bewertet, wie gut der durch KI generierte Text überprüft

wurde. Verständlich, dass wir meinen, ein durch KI generierter Essay verdiene weniger Anerkennung.

Doch nicht nur die Leser innen, sondern die Verfasserin oder der Verfasser selbst kann das Gefühl haben, dass die Arbeit weniger wert ist und kein Lob verdient. Es entsteht eine «achievement gap»: Die Arbeit verliert an Sinn für das Individuum, wenn die Automatisierung ihm die Möglichkeit nimmt, besondere Leistungen zu erreichen. Hier könnte man einwenden, dass es zweifelhaft ist, ob Studierende schriftliche Arbeiten als sonderlich sinnstiftend wahrnehmen. Was ihnen jedoch durch das Nutzen von ChatGPT verwehrt bleibt, ist die Erfahrung, wie sich die eigenen Fähigkeiten beim wissenschaftlichen Schreiben verbessern. Die Unsicherheiten, die die ersten Arbeiten im Studium begleiten können, zu überwinden und zu erfahren, wie sich mit der Zeit die eigenen Fähigkeiten verbessern, kann durchaus als wertvoll wahrgenommen werden.

Wenn wir also tatsächlich weniger Anerkennung für mit KI generierte Texte verdienen, bleibt die Frage: Was heisst das für die Praxis? Komplett auf schriftliche Arbeiten zu verzichten, scheint zumindest für manche Fächer kaum vorstellbar. Hingegen sollten wir Arbeiten inhaltlich und formal so anpassen, dass ChatGPT wenig sinnvoll genutzt werden kann. Das kann heissen, dass sich die Aufgabenstellung stärker auf die im Seminar geführten Diskussionen bezieht und mehr Wert auf kritische Reflexion gelegt wird, statt auf die Wiedergabe von Inhalten. Ausserdem könnten schriftliche Arbeiten mit mündlichen Prüfungen kombiniert werden, wie der Vorstellung der Arbeit im Seminar. Das hätte nicht nur für die Dozierenden den Vorteil, dass erkennbar wird, ob die Arbeit selbst verfasst wurde, sondern fördert auch Präsentationsfähigkeiten und den gegenseitigen Austausch der Studierenden über ihre Arbeiten.

Unsere Expertin ► Sarah Köglsperger ist Diplomassistentin am Departement für Philosophie. sarah.koeglsperger@unifr.ch

Literatur

- ▶ Danaher, J., Nyholm, S. «Automation, work and the achievement gap,» *AI Ethics* 1 (2021), 227–237.
- ▶ Eshleman, A. S. «Worthy of Praise: Responsibility and Better- than-Minimally-Decent Agency,» in Shoemaker, D. and Tognazzini, N. (Hg.), Oxford Studies in Agency and Responsibility, Volume 2, (Oxford: Oxford University Press 2014), 216–242.
- Porsdam Mann, S. et al. «Generative AI entails a credit-blame asymmetry,» Nat Mach Intell 5 (2023), 472-475.

La conscience artificielle – entre Pygmalion et Pandore

Après l'intelligence artificielle, l'étape suivante sera-t-elle la conscience artificielle? Assistons-nous à l'émergence d'une forme de conscience synthétique? La complexification des réseaux de neurones artificiels pourrait-elle produire une forme de conscience de soi? Chérine Fahim

Connaissez-vous l'histoire de Pygmalion? Le mythique sculpteur grec créa une statue représentant une femme si douce, si parfaite à ses yeux, qu'il en tomba follement amoureux, s'imaginant qu'elle était un être de chair et de sang. Face à une telle détermination, les dieux n'eurent d'autre choix que de réaliser son vœu en dotant cette femme «de marbre» d'un sourire, de la parole... en lui donnant vie. La foi de Pygmalion était si puissante qu'elle lui permit de transcender les contraintes d'une réalité qui semblait éternellement fixée dans le marbre et d'obtenir ce qu'il espérait ardemment. La morale de l'histoire est une prophétie auto-réalisatrice: le simple fait de croire en quelque chose augmente les chances que cette dernière se réalise.

L'être humain est à un moment exceptionnel du développement et de l'implantation de l'intelligence artificielle. Des questions sur les chances et les risques liés la conscience, au transhumanisme, aux humanoïdes et au libre arbitre se posent dans presque tous les pays du monde, à l'échelle locale, nationale et internationale. Une culture mondiale unique est-elle en train d'émerger? La conscience de l'intelligence artificielle remplacera-t-elle la conscience humaine? Vers où allons-nous: vers la statue de Pygmalion devenue consciente ou vers la boîte de Pandore?

Qu'est-ce qui nous définit?

Sommes-nous nos mémoires, teintées par nos émotions, gravées dans nos corps, âme et esprit? Sommes-nous

défini-e-s par nos réseaux de neurones qui créent nos mémoires? Ces mémoires créent-elles notre conscience? Quels sont les mécaniques de la biologie humaine nécessaires pour qu'un cerveau puisse acquérir les mécaniques de la conscience cognitive, sociale et incarnée. Même avant notre naissance, notre cerveau est capable d'apprendre et de mémoriser. Si nous apprenons grâce à une mémoire qui associe les évènements et permet leur récupération, qu'est-ce qui nous différencie de ChatGPT? Et si nous dotons l'intelligence artificielle de réseaux de neurones, deviendra-t-elle consciente?

Les études de la conscience permettent de jeter un regard profondément humain sur le processus de la pensée et de l'action. Tentons une définition de l'indéfinissable: Etre consciente, c'est savoir synchroniser ses réseaux de neurones, connaître ses peurs, désirs, valeurs, croyances et intentions; c'est être capable d'introspection passée, présente et future, de prendre conscience des expériences viscérales et sensorielles significatives. La conscience, c'est la capacité cérébrale à connecter différents réseaux de neurones et les synchroniser en un tout fonctionnel pendant un court moment dans le temps. La mémoire de travail (mémoire à court terme dans le lobe frontal), par exemple, nous permet de conserver dans notre «œil de l'esprit» le contenu de notre conscience consciente, même en l'absence d'entrée sensorielle, car elle maintient une représentation active de l'information pendant une brève période.



Rappelez-vous d'un être cher décédé, gardez son image mentale dans votre lobe pariétal. Existe-t-il dans votre cerveau, êtes-vous conscient e de ces beaux moments passés en sa compagnie?

Mécanique complexe

Les fonctions cérébrales œuvrent par activité électrique (les ondes cérébrales) et chimique (les neurotransmetteurs), considérées comme des calculs effectués par les synapses dans des réseaux de neurones qui intègrent l'information. Ces fonctions peuvent se dérouler de manière consciente ou non consciente. Comparons cette activité synaptique entre l'Humain et l'intelligence artificielle. La synapse est une structure où les projections de deux neurones se rencontrent pour passer l'information. Des courants électriques descendent le long d'une branche neuronale jusqu'à ce qu'ils atteignent une synapse. Là, ils se mélangent à toutes les impulsions provenant des autres branches et s'additionnent pour former un signal unique. Lorsqu'elle est suffisamment forte, l'électricité déclenche la libération par le neurone des neurotransmetteurs qui dérivent vers la synapse d'un neurone voisin et provoquent à leur tour le déclenchement du neurone. Voici l'élément crucial: à chaque fois que cela se produit, la synapse est légèrement modifiée, en ce sens qu'elle nécessite moins (ou plus) d'énergie pour activer le neurone en aval. Revenons à la synapse artificielle. Comme une synapse biologique, elle se compose de deux films minces, composés de matériaux organiques flexibles, séparés par un mince espace contenant un électrolyte qui laisse passer les protons. L'ensemble du dispositif est contrôlé par un interrupteur principal: lorsqu'il est ouvert, le dispositif est en mode «lecture seule»; lorsqu'il est fermé, le dispositif est «inscriptible» et prêt à stocker des informations. Pour introduire des données, les chercheurs euses soumettent la couche supérieure du film à une petite tension qui libère un électron. Pour neutraliser sa charge, le film «vole» ensuite un ion hydrogène au film voisin inférieur. Cette réaction modifie sa conductivité. A l'instar des synapses biologiques, plus l'impulsion électrique initiale est forte ou longue, plus les ions hydrogène sont déplacés, ce qui correspond à une plus grande conductivité. Avec de l'entraînement, les chercheurs euses peuvent prédire avec un pour cent d'incertitude la tension nécessaire pour atteindre un état particulier (Yang & Xia, 2017). Ainsi la synapse artificielle fut!

Donc, l'intelligence artificielle ne cesse de reproduire les fonctions de l'intelligence humaines. Aujourd'hui, l'intelligence artificielle est capable de théorie de l'esprit (TOM en anglais ou *Theory of Mind*). Cette caractéristique de l'intelligence sociale nous permet de déduire la réalité intérieure de l'esprit d'une autre personne sur la base de son discours et de son comportement, ainsi que de

notre propre connaissance de la nature humaine incluant nos peurs, désirs et intentions. La TOM est aussi un ingrédient essentiel de l'empathie, du jugement moral, de la conscience de soi et peut-être même du libre arbitre (Cuzzolin et al., 2020). C'est vers trois mois après la naissance que l'enfant effectue ses premiers pas vers la découverte de la conscience en découvrant les neurones miroir et la capacité d'imiter l'autre dans les grimaces et le sourire ou en découvrant la permanence de l'objet avec le jeu du coucou et l'attention conjointe. L'enfant n'a pas, à la base, ce concept de conscience de soi. C'est par le biais de de ses interactions que la séparation entre son «soi» et le «soi» de l'autre se développe au fil du temps.

Ancré dans le réel

Un constat qui nous met face à une réalité. Le cerveau humain présente un paradoxe déroutant et stimulant: malgré une anatomie fixe caractérisée par sa connectivité, son répertoire fonctionnel est vaste, permettant l'action, la perception et la cognition, des processus à la fois infiniment complexes, intuitifs et simples. La résolution de ce paradoxe peut résider dans l'architecture de réseau du cerveau, qui organise les interactions locales pour faire face à diverses demandes environnementales – assurant l'adaptabilité, la robustesse, la résilience aux dommages,

L'intelligence artificielle ne cesse de reproduire les fonctions de l'intelligence humaines

la transmission efficace des messages et les diverses fonctionnalités d'une structure fixe. Assurant même l'activité en absence d'entrée sensorielle: notre capacité d'imagination, de rêves éveillés, de partir ailleurs tout en étant là physiquement! Le cerveau humain est doté d'une conscience réelle et non virtuelle. Etes-vous capable d'écouter, dans votre tête, une musique qui n'existe pas? Sans entrée sensorielle réelle, pouvez-vous rejouer la chanson que vous aimez bien dans vos réseaux de neurones?

Le siège de la conscience est le cortex cérébral avec ses labyrinthes. Une compréhension anatomique et physiologique de la structure et du fonctionnement du cerveau confirme qu'il est le «siège de l'esprit». En tant qu'organe d'apprentissage, de réflexion et de mémoire, le cerveau devient synonyme de ce qui définit le «soi» à travers l'existence de la conscience, de l'esprit. Ainsi, le cerveau est indéniablement associé à une gamme de concepts transcendants – l'âme, l'esprit et la conscience – qui se rapportent tous de manière fondamentale les uns aux autres à

la fois en raison de leur emplacement perçu dans le cerveau et de la manière dont chacun travaille finalement à définir la personne à laquelle appartient le cerveau.

Bienvenue dans les méandres du cerveau, l'abîme de la conscience, la spiritualité ou le royaume de l'âme. Des chercheurs euses ont présenté un nouveau modèle neuro-informatique du cerveau humain qui pourrait permettre de mieux comprendre la façon dont le cerveau développe des capacités cognitives complexes et de faire progresser la recherche sur l'intelligence artificielle neuronale. Le modèle décrit le développement neuronal sur trois niveaux hiérarchiques de traitement de l'information qui obéissent au développement du cerveau humain. Au premier plan, le niveau sensorimoteur explore comment l'activité interne du cerveau intègre des régularités statistiques à partir de la perception sous forme de schémas et les coordonne avec l'action. Le niveau cognitif examine quant à lui la façon dont le cerveau combine ces schémas de manière contextuelle. Enfin, le niveau conscient considère comment le cerveau se dissocie du monde extérieur et manipule les schémas appris, et stockés dans la mémoire, qui ne sont plus accessibles à la perception, comme dans l'exemple d'un être cher décédé.

Avons-nous réussi la «création» d'une intelligence dotée d'une cognition artificielle, une cognition socio-émotionnelle, et allons-nous réussir à «créer» une cognition incarnée artificielle? La cognition incarnée représente les interactions entre pensées, affects et comportements au sein d'un environnement physique. Elle nous permet de concevoir le fonctionnement de la pensée dans un schéma intégratif (corps-cerveau-environnement) visant à prendre

L'esprit n'est pas un processus cérébral qui observe le monde à l'aide de capteurs. Il est, au contraire, une interaction entre les processus d'un cerveau et le monde qui l'entoure

en compte la nature incarnée des connaissances que nous possédons sur le monde. L'esprit n'est pas un processus cérébral qui observe le monde à l'aide de capteurs. Il est, au contraire, une interaction entre les processus d'un cerveau et le monde qui l'entoure. Les types de problèmes pour lesquels les réseaux neuronaux profonds (DNN – Deep Neural Networks en anglais) excellent sont précisément ceux où

l'interaction est au premier plan. Les décisions prises par les DNN sont frustrantes et inexplicables, et ne peuvent être qualifiées de rationnelles (Thelen, 2000; Lee, 2022)

Ces modèles artificiels démontrent comment la convergence entre les neurosciences et l'intelligence artificielle met en évidence les mécanismes biologiques et les architectures cognitives qui peuvent alimenter le développement de la prochaine génération de systèmes d'intelligence artificielle consciente. Sur quels mécanismes biologiques ces modèles neuro-computationnels s'appuient-ils? Sur l'interaction entre deux types fondamentaux d'apprentissage: l'apprentissage hebbien par utilisation et répétition et l'apprentissage par renforcement, associé à la récompense

Les futurs modèles de l'intelligence artificielle feront donc un pont entre l'intelligence artificielle et le seul système connu doté d'une conscience incarnée: l'être humain (corps-âme-esprit)

et aux neurotransmetteurs. Le modèle résout trois tâches de complexité croissante à tous ces niveaux, de la reconnaissance visuelle à la manipulation cognitive des perceptions conscientes. Ce modèle neuro-computationnel utilise l'apprentissage pour progresser davantage. Les résultats mettent en lumière deux mécanismes fondamentaux pour le développement à plusieurs niveaux des capacités cognitives dans les réseaux neuronaux: 1) l'épigenèse synaptique ou l'apprentissage hebbien à l'échelle locale et l'apprentissage par renforcement à l'échelle globale; 2) les dynamiques auto-organisées, issues de l'activité spontanée et de l'équilibre entre les neurones excitateurs et inhibiteurs (Volzhenin et al., 2022). Les futurs modèles de l'intelligence artificielle feront donc un pont entre l'intelligence artificielle et le seul système connu doté d'une conscience incarnée: l'être humain (corps-âme-esprit).

Notre experte Chérine Fahim est docteure en sciences neurologiques et chargée d'enseignement au sein de l'Unité de psychologie clinique et de la santé du Département de psychologie, sous la direction de la Professeure Chantal Martin Soelch. cherine.fahimfahmy@unifr.ch



«Pics or it didn't happen» war gestern

Deepfakes sorgen dafür, dass wir Bildern und Videos immer weniger trauen können. Die Technologie kann politische Desinformation und kriminelle Handlungen befördern. Kein Grund zur Panik, sagt die Juristin Nula Frei – und setzt dabei auf unseren Verstand und eine starke Zivilgesellschaft. Matthias Fasel

Wolodimir Selenski ruft in einem Video sein Volk zur Kapitulation auf, ein betagter Mann hört seine Enkelin am Telefon flehen, er solle ihr sofort 30'000 Franken für eine überlebenswichtige Operation überweisen, eine Frau sieht auf einem Bild bei WhatsApp, wie ihr Mann Sex mit einer Unbekannten hat. Was diesen drei Fällen gemein ist: Sie haben sich so nie abgespielt – und doch waren sie genauso zu sehen oder zu hören. Es handelt sich um Deepfakes. Unter dieser Bezeichnung, die sich aus den englischen Begriffen Deep Learning und Fake zusammensetzt, versteht man täuschend echt wirkende Audio-, Bild- oder Videoaufnahmen, die von künstlicher Intelligenz erzeugt werden. Schon kurze Aufnahmen reichen, um basierend auf dieser Originaldatei einen Menschen alles Mögliche sagen und machen zu lassen.

«Vielfach Angstmacherei»

Eine gruselige Vorstellung, doch die Juristin Nula Frei, die derzeit an der Studie «Deepfakes und manipulierte Realitäten» mitarbeitet, relativiert: «Je länger ich mich mit dem Thema befasse, desto weniger gruselig finde ich es. In der öffentlichen Debatte schwingt vielfach Angstmacherei mit, es wird die menschliche Kapazität ignoriert, kritisch zu sein und Sachen zu hinterfragen.» Frei betrachtet es als Aufgabe der Wissenschaft, die Diskussion zu versachlichen, aufzuzeigen, dass die Deepfake-Technologie auch für sinnvolle Zwecke verwendet werden kann – etwa in Form attraktiver Visualisierungen im Bildungsbereich oder bei der Rekonstruktion von Tathergängen in der Kriminalitätsbekämpfung.

Doch egal wie sachlich darüber debattiert wird, es bleibt dabei: Deepfakes bergen Gefahren. Nula Frei hat einen juristischen und politikwissenschaftlichen Hintergrund, kennt also sowohl die Gefahren der politischen Desinformation als auch jene im Bereich des Verbrechens. Was bereitet ihr am meisten Sorgen? «Der Aspekt der Kriminalität. Der Enkeltrick etwa ist halt schon noch einmal perfider und erfolgversprechender, wenn die Stimme gefälscht werden kann. Problematisch ist zudem der ganze Bereich der geschlechtsspezifischen Gewalt.»

Pornografie allgegenwärtig

Tatsächlich ist das Thema Deepfakes eng verbunden mit dem Thema Pornografie. «Die ersten Deepfakes waren pornografische Darstellungen, in denen die Gesichter prominenter Frauen in Sexszenen eingefügt wurden.» Auch heute ist die Problematik allgegenwärtig. Schätzungen zufolge machen pornografische Darstellungen 96 Prozent aller Deepfakes aus. «Solche Bilder oder Videos können auf zwischenmenschlicher Ebene grossen Schaden anrichten. Problematisch ist ausserdem, dass sie juristisch teils schwierig zu handhaben sind.»

Verwendet jemand unsere Stimme oder unser Bild ohne Einwilligung, ist das eine Persönlichkeitsverletzung, gegen die wir zivilrechtlich vorgehen können. Auf strafrechtlicher Ebene präsentiert sich die Situation jedoch komplexer. Es ist zwar strafbar, jemanden ungefragt pornografisch zur Schau zu stellen. Ein geschlossenes Forum im Internet etwa, gilt allerdings nicht als öffentliche Darstellung. Entsprechend fallen pornografische Deepfakes oft nicht unter den Tatbestand der Pornografie. Eine wichtige Ausnahme gibt es jedoch: Die Darstellung von Kindern bei sexuellen Handlungen ist immer strafbar – selbst wenn die Dateien nur auf dem eigenen Handy gespeichert werden. Wenn also Jugendliche von anderen Jugendlichen, die noch nicht 18 Jahre als sind, einen pornografischen Deepfake erstellen, ist das strafrechtlich automatisch relevant.

Braucht es Anpassungen im Strafrecht?

Und doch stellt sich die Frage: Bietet die aktuelle Schweizer Gesetzgebung noch genügend Schutz? Oder braucht es Anpassungen? «Es lohnt sich, darüber zu diskutieren, ob gewisse Aspekte nicht strafbar werden müssten», sagt Frei.

«Von jemandem ein Nacktbild zu erstellen, ist im Moment nicht strafbar, egal ob es weltweit verbreitet oder nur in einem Chat geteilt wird.» Der Grund: Ein Nacktfoto fällt nicht unter Pornografie, solange es bloss unangenehm und peinlich für die betroffene Person ist. Es muss schon eine klare sexuelle Konnotation aufweisen, also zum Beispiel eine Person in aufreizender Pose zeigen, um strafbar zu sein. So kommt es, dass es eine beliebte Spielart bei Deepfakes ist, Leute auszuziehen. «Da ist die Frage berechtigt, ob das nicht strafbar sein sollte. Denn sobald ein Handeln strafrechtlich relevant ist, liegt es an der Staatsanwaltschaft und der Polizei, Untersuchungen durchzuführen. Betroffene Personen müssen dann nicht mehr den zivilrechtlichen Weg beschreiten, bei dem die Beweislast bei ihnen liegt, sondern eine Anzeige genügt. Das ist eine Erleichterung», erklärt Nula Frei.

Probleme bei der Rechtsdurchsetzung

Die aktuelle Gesetzgebung ist aus Sicht der Juristin nicht das Hauptproblem. «Das grösste Problem ist die Rechtsdurchsetzung.» Was nützt es, nach einer Ehr- oder Persönlichkeitsverletzung auf dem beschwerlichen zivilrechtlichen Weg eine Unterlassungsklage einzureichen, wenn die Bilder in Sozialen Netzwerken bereits tausendfach geteilt wurden und niemand weiss, wer sie hochgeladen hat?

«Es braucht einerseits eine bessere Kooperation zwischen den Staaten, dass zum Beispiel eine Strafe auch in den USA durchgesetzt wird, wenn die Person in der Schweiz verurteilt wird», sagt Frei. «Andererseits braucht es mehr Pflichten für die grossen Plattformbetreiber, mit den Behörden zu kooperieren, ihnen beispielsweise IP-Adressen und die Identität anonymer Nutzer preiszugeben.» Bestrebungen sind im Gang, im August ist in der Europäischen Union das Gesetz über digitale Dienste in Kraft getreten. «Auch die Schweiz versucht, Plattformen besser zu regulieren, allerdings steht sie noch ganz am Anfang. Der erste Gesetzesentwurf soll 2024 folgen.»

Gefahr für die Demokratie?

Juristisch also wird nach Lösungen gesucht. Doch welche Auswirkungen hat es aus gesellschaftlicher und politischer Sicht, wenn wir unseren Augen und Ohren nicht mehr trauen können? Politiker_innen kann in Zukunft jede beliebige Aussage in den Mund gelegt werden. Darauf machte die News-Seite Buzzfeed schon vor fünf Jahren aufmerksam, als sie ein Deepfake-Video von Barack Obama erstellte, der darin Donald Trump beschimpfte. Es ist eines der bekanntesten Deepfake-Videos und war als Warnung gedacht, Buzzfeed wollte explizit auf die Gefahren für die Demokratie hinweisen. Nula Frei sieht es nicht ganz so dramatisch, sie vertritt den Standpunkt, dass durch Deepfakes nicht ein neues Problem entstanden ist, sondern ein bereits bestehendes Problem um eine weitere Ebene erweitert wurde. «Politische Desinformation ist kein neues

Phänomen. Fake News gibt es schon lange, sie werden mit solchen Bildern und Videos höchstens glaubwürdiger. Wir sollten keine Panik schieben, sondern weiter unser menschliches Urteilsvermögen stärken.»

Mediendidaktik an Schulen

Dabei ist ein Umdenken notwendig. «Pics or it didn't happen», lautet eine auf Social Media beliebte Aufforderung, wenn jemand mit etwas prahlt, ohne Beweise zu liefern. Doch nur weil etwas auf einem Bild zu sehen ist, heisst das längst nicht mehr, dass es auch passiert ist. Ein Beispiel dafür war das Foto von Papst Franziskus in Designer-Daunenjacke, das im Frühling für Aufsehen sorgte. «Es gab Leute, die das für echt hielten, es stellte sich dann aber schnell heraus, dass es ein Deepfake ist. So wird sich das entwickeln, die klassischen Medien, die mit ihrer Sorgfaltspflicht Informationen überprüfen und eine kritische Zivilgesellschaft werden schnell mit dem Finger darauf zeigen und Fakes entlarven», ist Frei überzeugt.

Wie aber schafft man es, dass möglichst viele Menschen dieser kritischen Zivilgesellschaft angehören und auch Bilder und Videos nicht immer für bare Münze nehmen? «Das Wichtigste ist die Stärkung der Medienkompetenz, und zwar vor allem im Bereich social media literacy. Mit Mediendidaktik und -erziehung in Schulen sowie Aufklärungskampagnen für die Bevölkerung sollte die kritische Haltung gefördert werden», sagt Frei. «Noch ist es anspruchsvoll, Deepfakes zu erstellen, die wirklich überzeugend sind. Das gelingt nur ressourcenstarken Akteur_innen. Es ist jedoch davon auszugehen, dass sich die Technik weiterentwickelt und die Schwelle niedriger wird. Umso mehr muss den Leuten klargemacht werden, dass es mehr braucht als Augen und Ohren.»

Die Technologie wird sich auch auf der Gegenseite weiterentwickeln. Detektionssoftware, die Deepfakes entlarvt, wird ebenfalls ausgeklügelter werden. Doch womöglich kann auch die irgendwann wieder umgangen werden. In diesem Katz-und-Maus-Spiel wird deshalb etwas immer wichtiger werden, das mit Technologie nichts zu tun hat: unser Verstand.

Matthias Fasel ist Gesellschaftswissenschaftler, Sportredaktor bei den «Freiburger Nachrichten» und freischaffender Journalist.

Unsere Expertin ► Nula Frei ist Lehr- und Forschungsrätin am Departement für internationales Recht und Handelsrecht. Sie hat Rechts- und Politikwissenschaft studiert und arbeitet derzeit an der von TA-Swiss in Auftrag gegebenen Studie «Deepfakes und manipulierte Realitäten» mit, die im Frühjahr 2024 publiziert wird. nula.frei@unifr.ch







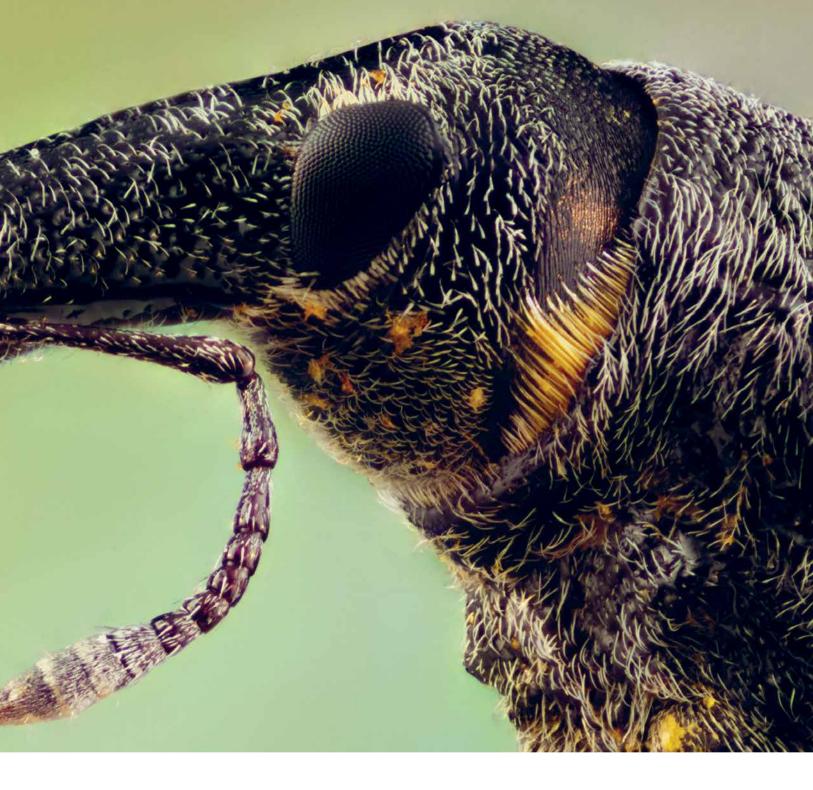
Der Weltbiodiversitätsrat IPBES hat kürzlich einen Bericht zu invasiven Arten veröffentlicht. Sie waren an der Erstellung dieses Berichts beteiligt. Was sind die Hauptaussagen?

Sven Bacher: Wir haben für den Bericht in vier Jahren Arbeit das aktuelle Wissen zusammengetragen. Der Bericht ist mehr als tausend Seiten lang. Zusammenfassend stehen drei Punkte im Vordergrund. Erstens: Invasive Arten sind ein riesiges und bisher weitgehend unbeachtetes und unterschätztes Problem. Zweitens: Das Problem ist

weltweit massiv am Wachsen. Und drittens: Eigentlich wissen wir, was wir tun müssten, um gegen das Problem vorzugehen. Aber leider tun wir es nicht, denn sonst wäre das Problem nicht am Wachsen.

Wieso wird das Problem unterschätzt?

Im Vergleich zum Klimawandel, der in aller Munde ist und von den meisten Personen um uns herum als wichtigste Bedrohung wahrgenommen wird, erhält die Biodiversitätskrise viel weniger Aufmerksamkeit. Das hat auch damit zu tun, dass wir dieses Phänomen nur schlecht fassen können. Wir haben zwar eine vage Vorstellung, ein Bauchgefühl, das uns sagt, dass wir die Natur schützen und dafür sorgen sollten, dass sie nicht leidet. Aber was genau meinen wir damit? Und wie – mit welchen Indikatoren – messen wir, wie gut es der Natur geht? Die einfachste und naheliegendste Antwort ist der Artenreichtum. Aber es genügt nicht, nur die Arten zu zählen. Denn wir müssen unterscheiden zwischen einheimischen Arten, die sich in Lebensgemeinschaften über lange Zeit



aneinander angepasst haben, und gebietsfremden Arten, die wir Menschen verfrachtet und irgendwo, bewusst oder unbewusst, neu angesiedelt haben. Einige dieser gebietsfremden Arten können die einheimischen Lebensgemeinschaften empfindlich stören, auch wenn sie kurzfristig zu einer Erhöhung der Artenzahl vor Ort beitragen. Auch diese Widersprüchlichkeit trägt dazu bei, dass wir – in der allgemeinen Bevölkerung, aber auch in der Wissenschaft – dem Problem der invasiven Arten wenig Beachtung schenken.

Wie viele der gebietsfremden Arten sind invasiv?

Von den weltweit etwas mehr als 37 000 gebietsfremden Arten, deren Ausbreitung wir im Bericht dokumentieren, verhalten sich 90 Prozent unauffällig – und sind also unproblematisch. Nur bei etwa 10 Prozent oder 3500 Arten haben wir Nachweise, dass sie Probleme verursachen. Etwa indem sie sich ausbreiten, oft auf Kosten von einheimischen Arten. Im Bericht verwenden wir den englischen Begriff «impact», der sehr breit gefasst ist und eine Veränderung eines

Systems bezeichnet. Das Spektrum reicht von geringen Beeinträchtigungen einzelner Individuen, zum Beispiel wenn Bäume wegen der Konkurrenz von invasiven Pflanzen langsamer wachsen und weniger Samen produzieren, bis zum lokalen oder sogar globalen Aussterben von einheimischen Arten. Aber im Bericht gehen wir auch auf Veränderungen ein, die sich vor allem auf Menschen auswirken. Das ist zum Beispiel bei verschleppten Landwirtschaftsschädlingen der Fall. Oder auch bei der Wasserhyazinthe, einer invasiven Schwimmpflanze,

die in Afrika die Oberfläche ganzer Seen überwuchert – und es den lokalen Fischern verunmöglicht, ihre Netze auszuwerfen.

Im Bericht steht an prominenter Stelle, dass für dessen Erstellung die enge Zusammenarbeit mit Vertreterinnen und Vertretern von Indigenen sehr wichtig war. Ja, das kennt man als Europäer weniger, weil auf unserem Kontinent nur noch im hohen Norden kleine Restpopulationen von indigenen Gemeinschaften zu finden sind. Doch in anderen Bereichen der Welt gibt es viel mehr Leute, die von und mit der Natur leben. Diese Personen merken auch rascher als wir, wenn sich Probleme in der Natur entwickeln. Und aus diesem Erfahrungswissen heraus haben viele von ihnen auch schon verschiedene Strategien entwickelt, wie man mit invasiven Arten umgehen kann. Deshalb ist es sehr wertvoll, diese Leute einzubeziehen.

Der Bericht stellt fest, dass invasive Arten weltweit Unheil anrichten – und dass in Afrika am wenigsten Auswirkungen von invasiven Arten nachgewiesen sind.

Ja, aber aufgepasst: Das heisst nicht, dass invasive Arten in Afrika weniger Schäden verursachen als in anderen Regionen der Welt. Sondern, dass wir weniger gut informiert sind, weil wir für Afrika weniger Daten zur Verfügung haben als etwa für Europa oder Nordamerika. Dieses Ungleichgewicht in den Daten hat auch bei der Endverhandlung der Zusammenfassung unseres Berichts für politische Entscheidungsträger - der so genannten «summary for policy makers» grosse Bedenken ausgelöst, weil sich viele afrikanische Staaten unterrepräsentiert und übergangen fühlten. Umso mehr, als invasive Arten für viele Menschen in Afrika tatsächlich ein grosses Problem darstellen, allein schon für die Nahrungsversorgung, denn dort wird mehr lokal konsumiert als bei uns. In Europa ist das Problembewusstsein in der Bevölkerung viel weniger weit entwickelt, weil wir Landwirtschaftsschädlinge einfach mit Pestiziden bekämpfen. Wir haben hier in der Schweiz auch weniger Schwierigkeiten mit invasiven Arten. Wir sorgen uns vielleicht,

dass sich die Tigermücke bei uns bald auch nördlich der Alpen etabliert. Aber die Ausbreitung der Tigermücke wird von Politik und Behörden meist isoliert betrachtet – und nicht als Teil eines viel grösseren und übergreifenden Problems verstanden. Und vor allem fehlt in unseren Breitengraden auch das Verständnis, dass sich unser Kaufund Reiseverhalten in anderen Regionen der Welt negativ auswirkt.

Was spielt bei der Verschleppung von Arten eine grössere Rolle: der Tourismus oder der Warenhandel?

Die Antwort hängt davon ab, auf welche Gruppe von Lebewesen man fokussiert. So werden etwa invasive Pflanzen oft von Touristen verschleppt. Wer zum Beispiel in Patagonien wandern geht, hat an den Schuhen meist noch Schmutz kleben, der von Schweizer Bergen stammt. In diesem Schmutz stecken Pflanzensamen: Sie würden staunen, wie viele verschiedene Pflanzen sich aus dem Dreck in den Schuhprofilen züchten lassen. Bei Tieren hingegen kommen andere Mechanismen zum Zug. Hier spielt der Internethandel eine grosse Rolle, der es Personen ermöglicht, sich - anonym und komplett unreguliert - auch exotische Haustiere zuzulegen. Wer möchte, kann Ameisen kaufen. Und zwar auch Ameisenarten, die als invasive Arten bekannt sind. Und die nachgewiesenermassen schon starke Schäden angerichtet haben, etwa im Mittelmeerraum, aber auch in den Tropen.

Dass gebietsfremde Ameisen Schäden verursachen, kann ich mir gut vorstellen. Aber mich hat erstaunt, dass in der Top-10-Liste der Arten mit dem grössten «impact» auch Füchse und Ziegen vorkommen.

Das sind Arten, die schon vor 150 bis 200 Jahren verschleppt wurden. Da war die Menschheit noch anders drauf. Damals galten Kolonialismus und Sklavenhandel als etwas Gutes. In Europa gab es Gesellschaften, deren Zweck darin bestand, Tiere wie Ziegen, Schweine, Hühner, Katzen, Igel und Füchse in die Kolonien zu exportieren, damit die Siedler sich auch in der neuen, exotischen Umgebung an den Tieren erfreuen konnten, die sie schon von daheim

kannten. In Australien sind deswegen zahlreiche Beuteltiere ausgestorben. Oft haben Seefahrer auch Ziegen auf kleinen Inseln ausgesetzt, damit sie auf der Rückreise wieder anlegen, einige Tiere jagen und sich mit Frischfleisch versorgen konnten. Aber Ziegen haben die Eigenschaft, Pflanzen bis zum Boden wegzufressen, so dass nur noch die Wurzeln übrig bleiben. Zahlreiche Pflanzen in Europa haben sich an diese Art des Frasses angepasst und können aus den Wurzeln heraus wieder neue Triebe bilden. Doch die einzigartige Vegetation, die sich zum Beispiel auf isolierten Inseln wie Galapagos entwickelt hat, verträgt diesen Frass gar nicht gut, dort haben die Ziegen riesige Schäden verursacht.

Im Bericht steht, dass invasive Arten an mehr als der Hälfte der Fälle von weltweitem Aussterben schuld oder zumindest beteiligt sind. Das ist eine unerwartet grosse Zahl.

In der öffentlichen Wahrnehmung galten bisher Landnutzungsänderungen als grösster Verursacher des Artensterbens. Die Hauptbotschaft des vor vier Jahren erschienenen Weltbiodiversitätsberichts war: Eine Million Arten sind vom Aussterben bedroht. Das tönt dramatisch. Aber erstens geht es hier um Arten, die noch nicht ausgestorben sind. Und zweitens ist diese Zahl eine grobe Schätzung. Wenn wir uns auf die harten Fakten konzentrieren und uns auf die Arten beschränken, die nachgewiesenermassen ausgestorben sind, wie wir das in unserem Bericht gemacht haben, dann sind etwas mehr als tausend Arten verschwunden. Unter diesen dokumentierten Fällen von Artensterben sind tatsächlich 60 Prozent aufgrund von invasiven Arten ausgestorben. Für mich zeigt die grosse Diskrepanz zwischen den tausend dokumentierten Fällen und der Million Arten, die vom Aussterben bedroht sind, dass es keinen Sinn ergibt, die verschiedenen Faktoren - also etwa Klimawandel, Landnutzungsänderungen, invasive Arten oder Umweltverschmutzung - gegeneinander auszuspielen. Denn eigentlich geht es nicht darum, welcher Faktor schlimmer und welcher weniger schlimm ist. Sondern darum, dass man dringend etwas tun muss, wenn man die Natur erhalten möchte.

Eigentlich lohnt es sich nur schon aus finanziellen Gründen, etwas gegen invasive Arten zu unternehmen. Denn sie verursachen riesige Kosten, wie der Bericht festhält.

Ja, und zwar in der Höhe von 423 Milliarden Franken – pro Jahr! Auch diese Zahl beruht nicht auf Schätzungen, sondern auf belegten Tatsachen. Weil nicht alle Schäden dokumentiert sind, und weil noch nicht alle Datenquellen in die Schadensdatenbank integriert sind, müssen wir sogar davon ausgehen, dass es zudem noch eine grosse Dunkelziffer gibt.

Wie setzen sich diese Kosten zusammen?

92 Prozent der Kosten sind direkte ökonomische Auswirkungen von invasiven Arten. Im Bericht bezeichnen wir sie als Schadenkosten. Darunter fallen etwa Ernteeinbussen oder die Beschädigung der Infrastruktur. Nur acht Prozent sind Massnahmenkosten. Sie umfassen den finanziellen Aufwand, der betrieben wird, um invasive Arten zu bekämpfen und zurückzudrängen.

Wie sind diese Kosten global aufgeteilt?

Der grösste Teil der Kosten entfällt auf die industrialisierten Länder in Europa und Nordamerika. Wie bei den Belegen für die Auswirkungen von invasiven Arten war es auch bei den Kosten schwierig, für zahlreiche Entwicklungsländer verlässliche Zahlen zu finden. Wenn man die Summen weltweit zusammennimmt, verursachen invasive Arten Folgekosten, die sich mit denen von grossen Naturkatastrophen wie Erdbeben oder Überschwemmungen vergleichen lassen. Meine Kolleginnen und Kollegen, die das Kapitel über die finanziellen Auswirkungen verfasst haben, haben ausserdem gezeigt, dass die Schadensumme dramatisch steigt: Sie hat sich in den letzten 30 Jahren alle zehn Jahre vervierfacht.

Hat das damit zu tun, dass in diesem Zeitraum auch die Reisetätigkeit und die globalen Warenflüsse zugenommen haben?

Ja, das ist so. Wir sehen in unseren Daten überall den gleichen Entwicklungstrend, bei Pflanzen genauso wie bei Tieren oder Krankheitserregern. Wir begegnen nicht nur den schon bekannten gebietsfremden

Arten, die sich immer weiter ausbreiten, sondern vor allem immer öfter Arten, die neu gebietsfremd sind. Für uns ist das ein klares Zeichen dafür, dass Lebewesen in zunehmendem Masse verschleppt werden.

Was kann man gegen die Ausbreitung invasiver Arten tun?

Am wirksamsten sind Präventionsmassnahmen. Internationale Gesetze schreiben
zum Beispiel vor, dass Pflanzensendungen
vor dem Versand auf das Vorkommen von
Schädlingen kontrolliert werden müssen.
Das wird aber leider nur bei ein bis zwei
Prozent der Sendungen gemacht. Die Sendungen könnten auch sterilisiert werden,
die Methoden hierfür sind längst etabliert.
Aber weil sie kosten, besteht ein Anreiz, solche Hygienemassnahmen zu umgehen.

Wenn eine invasive Art früh genug in einem neuen Gebiet entdeckt wird, kann sie häufig noch erfolgreich ausgerottet werden. Doch wenn sie sich am neuen Ort schon festgesetzt und ausgebreitet hat, wird es schwieriger. Dann kann man häufig nur noch versuchen, sie an der weiteren Ausbreitung zu hindern. Konventionelle Methoden wie etwa das Versprühen von Pestiziden sind teuer, mit einem sehr grossen Aufwand verbunden und gehen zudem mit Nebenwirkungen einher. Oft kommen nur noch biologische Agenten als Lösung in Frage.

Wie bitte: Invasive Arten zurückdrängen, indem man weitere gebietsfremde Arten aussetzt?

Ja, ich weiss, wegen der Aga-Kröte, die vor allem in Australien grosse Probleme verursacht, hat diese Methode einen schlechten Leumund. Doch die Kröte wurde vor über 80 Jahren zur Bekämpfung von Zuckerrohrschädlingen angesiedelt, zu einer Zeit, als man nicht viel über biologische Bekämpfung wusste - und einfach schaute, was nach der Freilassung passierte. Seither ist allerdings viel passiert. Heute sind Freilassungen gesetzlich geregelt. Und es ist vorgeschrieben, dass biologische Agenten vor der Freilassung umfassend getestet werden. Es gibt zahlreiche Erfolgsgeschichten. So stammt die vorhin erwähnte Wasserhyazinthe, die in Afrika ganze Seen überwuchert, ursprünglich aus

Südamerika. Dort lebt auch ein kleiner Rüsselkäfer, der sich auf die Schwimmpflanze spezialisiert hat und nichts anderes frisst. Dieser Käfer hat sich im Viktoriasee als spektakuläre Lösung erwiesen und die invasive Pflanze so weit zurückgedrängt, dass sie überhaupt keine Probleme mehr verursacht und die Fischer wieder normal ihrem Handwerk nachgehen können.

Ori Schipper ist freischaffender Wissenschaftsjournalist.

Was ist die IPBES?

Die Intergovernmental Science-Policy Platform on Biodiversity and Ecosystem (IPBES) – auch Weltbiodiversitätsrat genannt – ist eine UN-Organisation, die 2012 ins Leben gerufen wurde.

Die IPBES gilt als die wichtigste unabhängige Quelle für fundierte multidisziplinäre Wissenschaft über Biodiversität und Ökosystemleistungen. Sie hat das Ziel. Wissenschaft. Politik und Gesellschaft über den Zustand und die Entwicklung der biologischen Vielfalt und deren Bedeutung für die Menschheit zu informieren, den Bedarf an Kapazitäten und Aktivitäten zu priorisieren sowie die Gewinnung relevanten Wissens zu unterstützen. Zur IPBES gehören Mitgliedstaaten und Beobachter wie etwa UNO-Institutionen, Forschungseinrichtungen oder Universitäten. IPBES übernimmt eine ähnliche Rolle wie der Weltklimarat (Intergovernmental Panel on Climate Change IPCC).

Aktuell sind 139 Staaten Mitglied von IPBES, darunter auch die Schweiz.

⊳ www.ipbes.net/ias (Bericht) ■



Unser Experte ➤ Sven Bacher ist Professor für angewandte Ökologie an der Universität Freiburg. Er hat als koordinierender Leitautor des Kapitels über die Auswirkungen von invasiven Arten am neuen Bericht des Weltbiodiversitätsrats IPBES mitgeschrieben. sven.bacher@unifr.ch





Juliane Britz et son équipe viennent de démontrer que notre cerveau «voit» mieux dans le moment en suspend entre deux battements de cœur et à l'inspiration qu'à l'expiration. Un renversement de perspective qui remet les signaux corporels au centre de l'attention des neuroscientifiques et, surtout, la démonstration qu'en matière de perception visuelle, tout ne tourne pas autour du cerveau! Christian Doninelli

Comme toutes et tous les spécialistes, les neuroscientifiques ont parfois le nez dans le guidon et manquent du recul nécessaire pour étudier leur matière de prédilection dans sa globalité. «Nous sommes des snobs encéphalocentrés», confesse Juliane Britz, lectrice au Département de psychologie de l'Université de Fribourg. Plutôt que de garder son regard exclusivement rivé sur les circonvolutions de notre encéphale, cette merveille qui gouverne chacun de nos gestes et de nos pensées, la chercheuse a décidé de «réamarrer le cerveau à notre corps», afin de voir si ce dernier joue également un rôle dans notre capacité à percevoir un stimulus visuel très faible. En 2020, elle a démarré une recherche s'inscrivant dans le cadre du projet «BBC -

La chercheuse a décidé de «réamarrer le cerveau à notre corps»

Brain, Body and Consciousness», soutenu par le Fonds national de la recherche scientifique et doté d'un financement de près de 400'000 francs. Un sujet qui a passionné les spectateurs-trices et le jury fribourgeois de Ma thèse en 180 secondes, puisque sa doctorante Viviana Leupin a gagné sa place en finale suisse en 2023 avec sa thèse «Comprendre l'interaction entre les signaux

cardiorespiratoires et l'activité cérébrale dans l'émergence de la conscience».

Renversement de perspectives

Cela fait des décennies que le cerveau fascine Juliane Britz, cet organe «toujours actif, jour et nuit, même sous anesthésie générale ou dans le coma, jusqu'à la mort». Après une thèse sur le traitement du langage chez les cérébrolésé·e·s, la chercheuse s'avoue toutefois un peu lassée des expériences à sens unique propres aux neurosciences cognitives. «Nous nous sommes trop souvent bornés à observer les effets des stimuli sur le cerveau. J'ai décidé de renverser la perspective et de centrer mes recherches sur ce que ce dernier fait des stimulations sensorielles qui lui parviennent.» Pour illustrer son propos, Juliane Britz brandit une image du vase de Rubin, une image en noir et blanc qui peut tantôt être perçue comme un vase, tantôt comme deux visages qui se font face. «Ce sont des images ambiguës que nous ne sommes pas capables de voir en même temps. C'est soit l'une, soit l'autre, ce qui démontre que, pour un même stimulus, le cerveau peut décider de distinguer deux formes différentes».

Nous avons toutes et tous entendu parler de ces images subliminales de quelques millisecondes que des publicitaires auraient intercalées sur les pellicules projetées au cinéma. Si l'efficacité marketing de l'injonction «Buvez Coca Cola!», glissée subrepticement entre l'une des 24 images à la seconde d'un film – le cas le plus célèbre – n'a pas été démontrée, l'idée n'en était pas moins la suivante: une image présentée un temps extrêmement bref, si bref qu'elle n'est pas

«Nous nous sommes trop souvent bornés à observer les effets de stimuli sur le cerveau. J'ai décidé de renverser la perspective et de centrer mes recherches sur ce que ce dernier fait des stimulations sensorielles qui lui parviennent» Juliane Britz

consciemment perçue, peut être enregistrée par le cerveau des spectatrices et spectateurs sans qu'elles et ils ne s'en rendent compte. C'est peu ou prou la même stratégie qu'a mise en œuvre Juliane Britz: «Nous leur montrons très fugacement des images appelées *Gabor gratings*, qui consistent en des traits diagonaux dessinés sur un cercle. Les participant·e·s doivent déterminer si ces motifs sont orientés à gauche ou à

droite. Nous avons fait en sorte qu'elles et ils puissent trouver la réponse correcte dans plus de 80% des cas». L'apparition étant extrêmement brève (16 millisecondes), au seuil de la perception, cela implique que les sujets ne voient le stimulus que la moitié du temps. «Les participant·e·s doivent nous indiquer d'abord l'orientation du stimulus et ensuite si elles et ils ont vu le stimulus ou deviné la réponse. De notre côté, nous comparons la réaction du cerveau pour le même stimulus, qu'il ait été vu ou pas, séparément pour la phase cardiaque et respiratoire», explique Juliane Britz.

Une expérience façon Orange mécanique

Imaginez une salle plongée dans l'obscurité avec un·e volontaire, assis·e à 70 centimètres d'un écran, la tête enveloppée dans une sorte de bonnet de natation constellé d'électrodes. «Nous en installons 128, afin de mesurer l'activité (électrique) cérébrale (électroencéphalogramme). Pour la pulsation cardiaque, nous posons deux électrodes (électrocardiogramme). Quant à la respiration, nous la mesurons à l'aide d'une ceinture autour de la taille. Une fois ces préparatifs achevés, qui auront tout de même duré près de deux à trois heures, les tests peuvent commencer. C'est ce protocole élaboré qui a permis aux chercheuses de voir comment le cerveau, au cours de plus de 960 essais, réagit aux images présentées à différents moments de leur cycle cardiaque (systole/diastole) et de leur cycle respiratoire (inhalation/exhalation). Place maintenant aux résultats.

Le cœur conditionne le cerveau...

Premier constat: la phase cardiaque a un impact indéniable sur la conscience, d'où le parallèle avec la révolution copernicienne, toute proportion gardée bien entendu. En d'autres termes, les expériences des deux chercheuses ont montré que notre palpitant affecte la manière dont nous percevons les choses. Il a rendu certaines informations plus claires durant la phase de décontraction du muscle cardiaque (diastole) que lors de la contraction (systole). Durant la systole, les barorécepteurs, des récepteurs sensoriels présents dans les artères, envoient des signaux au cerveau afin de lui

permettre de correctement régler la tension artérielle. Or, les chercheuses ont remarqué que ces signaux interfèrent avec le traitement des stimuli visuels par le cerveau. «En somme, explicite Juliane Britz, les sujets voient le stimulus tant dans la phase

«Ces résultats fournissent une nouvelle solution à un long débat scientifique sur l'importance des cortex préfrontaux et pariétaux pour la conscience» Juliane Britz

diastolique que systolique. En revanche, le cerveau utilise d'autres voies pour parvenir à une perception consciente, lorsqu'il y a un signal concomitant des barorécepteurs». Concrètement, cela signifie que, en fonction du signal corporel dispensé au travers des phases cardiaques, l'activité cérébrale prend une voie différente. «Ces résultats fournissent une nouvelle solution à un long débat scientifique sur l'importance des cortex préfrontaux et pariétaux pour la conscience. Grâce à nos tests, nous avons montré que cela dépend des signaux du corps qui fluctuent à travers la phase cardiaque et respiratoire. Ce même modèle nous a aussi permis d'identifier les premiers marqueurs de la conscience: nous trouvons des premiers marqueurs plus précoces quand les barorécepteurs se taisent que quand ils sont actifs», ajoute la chercheuse.

... et la respiration aussi

Les chercheuses ont également démontré que la phase de la respiration, inhalation-exhalation, influence aussi la trajectoire de l'activité cérébrale liée à la conscience. «Les barorécepteurs sont plus actifs pendant l'exhalation que pendant l'inhalation. Nous observons donc des effets comparables autant lors de la diastole et de l'exhalation que de la systole et de l'inhalation», détaille la chercheuse. Une influence de la respiration

sur le cerveau et notre conscience qui ne la surprend d'ailleurs pas tellement puisque «la première chose que nous faisons à la naissance, c'est souffler».

Cette vaste étude a permis aux deux neuroscientifiques de confirmer que le cerveau n'encaisse pas passivement les stimuli externes, mais que c'est son état pré-stimulus, déterminé par l'influence du corps, qui va déterminer sa manière de les traiter. Et comme toute bonne recherche produit plus de questions que de réponses, Juliane Britz anticipe déjà la suite: «Je rêve maintenant de poursuivre le projet avec des singes et des patient·e·s épileptiques, ainsi que sur des hallucinations – des perceptions sans stimulus physique».

Christian Doninelli est rédacteur à Unicom.

Notre experte ➤ Juliane Britz est lectrice au Département de psychologie et chargée de cours au Département d'informatique.
juliane.britz@unifr.ch



étudiant·e·s dans la vie active et permettre aux employeuses et employeurs de dénicher leurs futurs talents. Pierre Jenny

Un rapport de l'Université de Fribourg daté de juin 2022 indique que le taux de chômage est près de quatre fois plus élevé chez les diplômé·e·s de l'Université de Fribourg ayant débuté leurs recherches d'emploi après leurs études que chez les personnes ayant commencé ces recherches durant leur cursus universitaire. Faciliter les relations et tisser des liens entre les actuel·le·s étudiant·e·s et les entrepreneurs·euses est donc capital selon Astrid Epiney, rectrice de l'Université de Fribourg, et Jerry Krattiger, directeur de la Promotion économique du Canton de Fribourg. Interview croisée.

Quel est le but de ce salon?

Astrid Epiney: L'Université est déjà en contact étroit avec la politique, l'économie et la société en général dans le Canton de Fribourg. Mais ces contacts concernent peu, jusqu'à maintenant, les relations directes entre nos étudiant·e·s, nos futur·e·s diplômé·e·s et le monde économique. Le 14 mars 2024, le Career Forum permettra précisément de donner la possibilité à ces deux mondes de se rencontrer et de tisser des liens sur le campus de Pérolles. Les futur·e·s diplômé·e·s pourront entrer en contact avec des employeurs·euses potentiel·le·s et ces derniers·ères rencontrer d'éventuel·le·s futur·e·s employé·e·s.

Jerry Krattiger: Nous nous trouvons aujourd'hui dans une guerre pour les talents. Les derniers chiffres du Secrétariat d'Etat à l'économie (SECO) indiquent que 40% des entreprises en Suisse ont de la peine à recruter. J'ai des retours de certaines entreprises, et non des moindres, qui m'informent qu'elles ont de la peine à trouver le point d'entrée pour rencontrer les futurs talents. Le Career Forum permettra aux entreprises d'avoir un accès privilégié aux étudiant·e·s non seulement de l'Université, mais aussi de la Haute école de gestion (HEG-FR) et de la Haute école d'ingénierie et d'architecture (HEIA-FR) qui prennent part à ce salon.

A vous entendre, les deux milieux ne se connaissent pas vraiment...

Astrid Epiney: Depuis quelques années déjà, l'Université de Fribourg a créé un service permettant de tisser des liens entre nos

étudiant·e·s, nos doctorant·e·s et post-doctorant·e·s et le monde du travail en général. Ces «Career Services» ont été renforcés l'an dernier mais leurs moyens restent malgré tout limités. Nous travaillons aussi avec les différentes facultés et l'Association générale des étudiants et étudiantes de l'Université de Fribourg (AGEF) qui, via leurs propres canaux et leurs sites internet, relaient des informations et notamment les offres d'emplois. J'ajoute que bon nombre de nos étudiant·e·s ont déjà un lien avec le monde du travail en effectuant des stages ou en exerçant une activité professionnelle durant leurs études. Enfin, si les étudiant·e·s fribourgeois·e·s connaissent bien les entreprises du canton, ce n'est pas forcément le cas des étudiant·e·s qui viennent d'ailleurs. Il faut rappeler que deux tiers de nos étudiant·e·s proviennent d'autres cantons, voire de l'étranger. Le Career Forum est donc une occasion de resserrer les liens existants et d'en créer de nouveaux, tout en présentant les opportunités professionnelles afin que les étudiant·e·s restent à Fribourg une fois leurs études achevées.

«Le Career Forum est une occasion de resserrer les liens existants et d'en créer de nouveaux»

Astrid Epiney

Jerry Krattiger: Il est clair qu'on peut étudier et vivre à Fribourg, mais qu'on peut également y travailler. Notre canton vit une croissance démographique extrêmement dynamique ces dernières années. En parallèle, le produit intérieur brut par emploi plein temps a, lui aussi, augmenté de manière constante et forte. L'économie cantonale est dynamique. Les investissements réalisés par l'Etat dans les infrastructures pour absorber la croissance démographique ont été bien maîtrisés. Fribourg jouit donc d'une situation enviable et dispose d'un très bon tremplin de développement, tant économique qu'humain. Actuellement, quelque 40'000 Fribourgeois·e·s sortentdes frontières cantonales chaque jour pour aller travailler. Dans le sens inverse, seules 17'000 personnes d'autres cantons viennent travailler à Fribourg. Il y a un potentiel existant localement qu'il faut mieux exploiter, notamment pour des questions de durabilité. Dès le moment où vous étudiez, vous vivez et vous travaillez localement, votre impact environnemental est moindre. Le fait de ne pas connaître les talents issus de l'Université et des hautes écoles ne permet pas aux entreprises locales de bénéficier pleinement de cet important potentiel humain.

Quels sont les profils recherchés par les entreprises?

Jerry Krattiger: Tout d'abord, il faut rappeler que le Canton de Fribourg a une stratégie de développement économique axée sur la bioéconomie et l'industrie 4.0. Evidemment, cela va influer sur les profils recherchés. Pour la bioéconomie, on pense notamment à l'agroalimentaire, les matériaux biosourcés dans la construction et les biotechnologies. Là, les personnes ayant étudié la chimie, la biochimie, la gestion de processus biotechnologiques, la science des matériaux, etc., sont particulièrement recherchées. Pour l'industrie 4.0, on privilégiera bien sûr les profils en ingénierie, informatique, robotisation, automation ou encore en digitalisation.

Ensuite, le monde de l'économie est toujours à la recherche de profils plus généralistes comme des économistes, des juristes, des politologues ou des historien·ne·s, car ces personnes ont plusieurs cordes à leur arc et sont prédisposées à travailler de manière transverse et holistique. Enfin, il existe des besoins plus spécifiques. Une personne ayant étudié la philosophie ou la théologie peut travailler dans une commission d'éthique, par exemple. Quant aux anthropologues, leurs connaissances peuvent servir à l'observation des comportements des consommateurs trices ou à décrypter comment un·e utilisateur·trice interagit avec un produit ou une technologie. Les méthodologies enseignées en anthropologie peuvent aussi être utilisées à des fins économiques.

Pour l'Université, quelles sont les entreprises intéressantes aux yeux du monde académique?

Astrid Epiney: Nous sommes ouverts à tous les types d'entreprises. Avec ses cinq facultés – bientôt six –, l'Université de Fribourg

certains profils de postes impliquent une formation plus précise, voire des profils plus disciplinaires, également présents au sein de nos facultés. Le Career Forum va permettre à nos étudiant-e-s de mieux cerner la diversité des emplois possibles.



Astrid Epiney est professeure de droit international, droit européen et droit public suisse. Après avoir été doyenne de la Faculté de droit, puis vice-rectrice, elle occupe le poste de rectrice de l'Université de Fribourg depuis 2015. Elle est également présidente de la Chambre des hautes écoles universitaires de swissuniversities. astrid.epiney@unifr.ch.

propose des profils professionnels très diversifiés. Notre force réside, entre autres, dans notre bilinguisme et dans notre large palette de disciplines qui peuvent satisfaire les différentes entreprises fribourgeoises, sans oublier les institutions étatiques et paraétatiques, comme l'Agroscope fédéral qui construit son nouveau siège à Posieux. De plus, de nombreuses places de travail ne nécessitent pas un profil trop typé. Si un e entrepreneur-euse recherche une personne dans l'administration, la gestion de projets, etc., il peut trouver parmi nos étudiant-e-s de nombreux profils qui sont à même de remplir de tels cahiers des charges. Mais, bien entendu,

Au-delà des emplois et de la recherche de talents, dans quels autres domaines fautil renforcer les liens entre l'économie et les milieux académiques?

Jerry Krattiger: Il y a tous les éléments liés au transfert technologique. Au niveau cantonal, nous constatons une augmentation du nombre de projets soutenus par l'agence suisse pour l'encouragement de l'innovation, Innosuisse. La tendance à voir des entreprises travailler en collaboration étroite avec des hautes écoles sur des projets innovants se renforce. Fribourg soutient aussi de tels projets au travers de la Loi cantonale sur la promotion économique et de la Nouvelle

Politique Régionale. C'est un domaine passionnant et les demandes de collaboration sont en augmentation. Dans ce cadre-là, les centres de compétences des hautes écoles jouent un rôle central. Les liens entre milieux économiques et académiques concernent aussi la commercialisation d'un projet de recherche. Cette réflexion est omniprésente dans le monde anglo-saxon. «Y a-t-il un moyen de générer une activité commerciale, de faire de l'argent avec un projet ou un savoir spécifique?» Cette question ne se pose pas automatiquement en Suisse, en France ou en Allemagne, même si je constate un changement culturel à ce sujet. J'en veux pour preuve la création de start-up intéressantes issues de l'Université de Fribourg, par exemple Neuria qui a reçu le Prix à l'innovation du Canton de Fribourg dans la catégorie Start-up. Elle a développé une thérapie digitale à travers une application de jeux vidéo qui permet de renforcer des habitudes alimentaires saines. Cette année, Fribourg place trois jeunes entreprises parmi le top 100 des start-up en Suisse. L'une d'elle, Impossible Materials, est issue de compétences développées aux Universités de Fribourg et de Cambridge. Ceci démontre que la dimension économique et entrepreneuriale d'un projet de recherche est tout à fait pertinente.

Astrid Epiney: La recherche appliquée est un aspect important. L'Université de Fribourg a créé un service spécialisé dans le transfert de connaissances et de technologies, le KTT. Ce service soutient les contacts entre nos chercheurs euses et l'économie, ainsi que toute entreprise désireuse de travailler avec nos chercheurs·euses. Mais la recherche appliquée ne doit pas occulter l'importance de la recherche fondamentale, conduite par la seule curiosité, sans objectif précis ou de commercialisation. Il y a de nombreux exemples où ce type de recherches a permis de trouver des produits auxquels personne n'avait pensé. Je crois qu'il est aussi important de garder une complémentarité avec les hautes écoles, les HES, qui sont focalisées presque uniquement sur la recherche appliquée. Nous avons près d'un millier de projets collaboratifs avec ces écoles. Cette complémentarité permet aussi de faire rayonner le Canton de Fribourg et d'augmenter son attractivité. Pour une entreprise qui désirerait s'installer à Fribourg, savoir qu'elle peut bénéficier des recherches et des activités menées à l'Université est un atout indéniable. Mais je tiens à rappeler que cette attractivité est liée à l'importance des infrastructures, d'ailleurs aussi cruciales pour des projets financés par des fonds tiers, qu'ils soient publics ou privés. Or, en ce qui concerne l'Université, il y a une réelle marge

mais n'oublions pas que tant la construction du nouveau bâtiment pour la Faculté de droit sur le site dit «Tour Henri» que les différents projets sur le plateau de Pérolles seront décisifs pour maintenir et développer l'attractivité de l'Université, non seulement pour les étudiantes et chercheurs euses, mais aussi pour l'économie fribourgeoise et les entreprises intéressées à s'installer dans le canton.



Jerry Krattiger est directeur de la Promotion économique du Canton de Fribourg depuis 2019. Il a auparavant été directeur national pendant plus de 4 ans de Youth For Understanding (YFU), menant la gestion opérationnelle, financière et stratégique de l'association. jerry.krattiger@fr.ch

de progression. L'avancement des différents projets d'infrastructures universitaires est crucial, car il prouve que le Canton de Fribourg investit dans ce domaine et renvoie une image positive, notamment auprès des milieux économiques. Je rappelle que le dernier gros investissement dans du bâti universitaire date de 2005 avec la création du site de Pérolles 21. Depuis, il y a certes eu le pavillon de médecine en 2015–2017,

Dans le monde anglo-saxon, les milieux économiques prennent souvent part au financement de chaires universitaires ou de projets de recherche. Ce modèle inspire-til l'Université de Fribourg?

Astrid Epiney: Le système suisse est très différent de celui des Etats-Unis ou de l'Angleterre. Les universités suisses, dont celle de Fribourg, sont très majoritairement financées par des fonds publics. A mon avis,

notre système a fait ses preuves, notamment sur le coût des études qui reste modeste en comparaison des taxes que vous devez payer aux Etats-Unis, par exemple. La qualité d'enseignement et de recherche de nos universités est également très élevée. Pour Fribourg, la part de financement privé représente 5% environ de tous les fonds tiers, les 95% provenant de fonds tiers publics, tels que le Fonds national. Nous avons des directives, publiée sur notre site internet, qui précisent les conditions et modalités d'un financement privé. Il faut que le domaine en question corresponde à la stratégie globale de développement de l'Université. L'autre aspect consiste à préserver notre liberté académique, y compris pour les engagements du personnel. Elle est le gage de notre réputation et de notre crédibilité. Un jour, une fondation nous a proposé un don important pour financer une chaire, mais en essayant de nous imposer également le professeur... Nous avons refusé et je suis convaincue qu'il est important de le faire.

Première édition du Career Forum en mars 2024: est-il déjà prévu de pérenniser cet événement?

Astrid Epiney: Nous allons d'abord tirer un bilan de cette première édition et, en fonction de celui-ci, nous verrons de quelle manière adapter l'événement et quelle fréquence lui donner. Mais, en effet, je pars de l'idée que ce genre d'évènement est appelé à avoir lieu régulièrement au sein de notre alma mater.

Jerry Krattiger: Un premier bilan est nécessaire. Mais une chose est certaine et c'est le message que j'aimerais faire passer au sein des milieux économiques fribourgeois: on ne peut pas dire, d'une part, qu'on a de la peine à trouver et à recruter des talents et, d'autre part, ne pas participer à un tel salon. C'est une chance de tisser des liens privilégiés avec nos hautes écoles et j'espère que le monde de l'économie saura saisir cette opportunité.

Pierre Jenny est journaliste indépendant.



Die Lehrpersonenausbildung im Kanton Freiburg soll an einer Institution vereint werden: Ab Herbst 2025 gehört die Pädagogische Hochschule (PH) mit der Ausbildung der Primarlehrpersonen zur neu geschaffenen Fakultät für Erziehungsund Bildungswissenschaften der Universität Freiburg. Jean-Luc Brülhart

Die Zusammenführung der Pädagogischen Hochschule (PH) mit der Universität Freiburg ist ein Leuchtturm-Projekt für den Kanton. Eines, das zudem schweizweit ausstrahlt, denn «wir spüren, dass wir mit der Schaffung einer neuen Fakultät in der Bildungslandschaft unter Beobachtung stehen und dass grosses Interesse an unserer Vorgehensweise besteht», sagt Floriane Gasser, Vorsteherin des Amtes für Universitätsfragen und Präsidentin des Projektausschusses. In der neuen Fakultät für Erziehungs- und Bildungswissenschaften wird die Primarlehrer_innenausbildung der PH mit den bestehenden Ausbildungen der Uni für die Sekundarstufe I und II sowie mit der Sonderpädagogik und den Erziehungswissenschaften vereint.

«Interessante Perspektiven»

Die organisatorische und legislative Vorarbeit erfordere ein grosses Engagement vieler Beteiligter, am Ende aber stehe «ein einmaliges Kompetenzzentrum rund um Erziehung und Bildung von Kindern und Jugendlichen», ist Astrid Epiney, Rektorin

«Durch die Zusammenführung der gesamten Lehrpersonenausbildung entstehen für Studierende und Dozierende viele Synergien»

Astrid Epiney

der Uni Freiburg, überzeugt. «Durch die Zusammenführung der gesamten Lehrpersonenausbildung entstehen für Studierende und Dozierende viele Synergien. Darüber hinaus werden sich für angehende Lehr-

personen durch den Einbezug von weiteren Disziplinen - insbesondere der Erziehungswissenschaften - interessante Perspektiven der Zusammenarbeit in Forschung und Lehre ergeben.» Diese Meinung teilt die Rektorin der Pädagogischen Hochschule, Delphine Etienne-Tomasini, und fügt hinzu: «Der Wissensaustausch zwischen den Primar- und den Sekundarstufen wird erleichtert und die Zusammenarbeit im Bereich der Weiterbildung gestärkt.» Nicht zuletzt soll die Einbettung der Primarlehrpersonen-Ausbildung in eine grössere Institution für mehr deutschsprachige Studierende sorgen. Im Herbst 2023 standen in der PH 28 deutschsprachigen Studienanfänger_innen 150 französischsprachige gegenüber.

Integration des PH-Personals

Neben der konzeptionellen Vorbereitung steht aber die Frage im Raum: Wie gehen die Mitarbeitenden der PH mit der Tatsache um, dass ihre Institution aufgehoben wird und sie in die Universität integriert werden? Dazu Delphine Etienne-Tomasini: «Das Projekt ist nun so weit gereift, dass wir den Studierenden sowie dem Personal konkrete Antworten liefern können.» Jedenfalls geniesst das gute Gelingen der Integration des PH-Personals in die universitären Strukturen von Seiten der Projektleitung höchste Priorität. «Dies ist zweifellos die grösste Herausforderung der Zusammenführung. In allen Diskussionen und von beiden Seiten habe ich vor allem Pragmatismus, gegenseitiges Verständnis und den Willen, möglichst zufriedenstellende Lösungen zu finden, gespürt», sagt Sylvie Bonvin-Sansonnens, Direktorin für Bildung und kulturelle Angelegenheiten und Präsidentin des Steuerungsausschusses. Floriane Gasser ergänzt: «Es geht nicht nur um die juristische Überführung, sondern viel mehr um den emotionalen Aspekt: Welche Funktionen werden definiert, welche Titel vergeben?» Dies gilt für die Dozierenden und Fachper«In allen Diskussionen und von beiden Seiten habe ich vor allem Pragmatismus, gegenseitiges Verständnis und den Willen, möglichst zufriedenstellende Lösungen zu finden, gespürt»

Sylvie Bonvin-Sansonnens

sonen ebenso wie für das Verwaltungs- und technische Personal. Die Zusammenlegung soll auf das Herbstsemester 2025 erfolgen – zur Unterstützung dieses Change-Managements wurde eine Arbeitspsychologin eingesetzt. Sie begleitet beide Institutionen auf diesem Weg.

Zwei Kulturen

Die Eingliederung der PH in die Uni ist auf die Harmonisierung der obligatorischen Schule und die Einführung der Lehrpläne für die West- und die Deutschschweiz (PER und Lehrplan21) zurückzuführen. Sie beruht auf einer pädagogischen Gesamtvision von der 1H bis zur 11H. Die bisherige Aufteilung der Lehrpersonenbildung auf zwei Institutionen - die PH für die Stufen 1H bis 8H, die Uni für die 9H bis zur 11H ist demnach nicht mehr zeitgemäss. «Dieses Projekt haben wir auf einem weissen Blatt Papier gestartet, aber sehr schnell im Prozess wurde allen klar, dass die Schaffung einer neuen universitären Fakultät die beste und zukunftsweisendste Lösung ist», sagt Floriane Gasser. Eine Arbeitsgruppe bestehend aus Mitgliedern der Uni und der PH ist aktuell daran, die genaue Struktur des neuen Departements Lehrpersonenbildung zu definieren. Es müssen organisatorische

Fragen geklärt werden, aber auch, wie den beiden Kulturen gerecht zu werden ist. Dazu Delphine Etienne-Tomasini: «Es liegt uns viel daran, dass ein Gefühl der institutionellen Zugehörigkeit entsteht und alle Mitarbeitenden der heutigen PH ihre Bezugspunkte und den Sinn im neuen beruflichen Kontext finden. Der Austausch mit den Kolleginnen und Kollegen der Uni wird helfen, eine gemeinsame Identität zu entwickeln.»

«Der Austausch mit den Kolleginnen und Kollegen der Uni wird helfen, eine gemeinsame Identität zu entwickeln»

Delphine Etienne-Tomasini

Eine wichtige Stelle ist schon besetzt. Anfang Oktober wurde Vivien Rüffieux zur administrativen Leiterin der neuen Fakultät ernannt. Vivien Rüffieux hat an der Universität Freiburg zweisprachig studiert und in Germanistischer Sprach- und Literaturwissenschaft promoviert. Sie verfügt zudem über einen Master of Arts in Rechtswissenschaftlichen Studien sowie die IPMA-Zertifizierung in Projektmanagement. Zuletzt war die 32-Jährige als wissenschaftliche Mitarbeiterin im Generalsekretariat der PH Bern tätig.

Vorerst existiert die neue Fakultät nur institutionell unter einem Dach. «Ein gemeinsamer Campus für die gesamte Fakultät ist und bleibt natürlich ein erklärtes Ziel. Die räumliche Nähe würde Synergien fördern und den Austausch unter Mitarbeitenden erleichtern», ist Floriane Gasser überzeugt. Wenn von einem Neubau die Rede ist, dann ist von einem Zeithorizont von mindestens 15 Jahren auszugehen.

«Ein gemeinsamer Campus für die gesamte Fakultät ist und bleibt natürlich ein erklärtes Ziel»

Floriane Gasser

Mittelfristig soll zumindest das Departement für Lehrpersonenbildung an einem Ort vereint werden. Noch ist keine Örtlichkeit spruch- respektive druckreif. Aber, so Floriane Gasser, «wir verfolgen momentan interessante Varianten.»

Jean-Luc Brülhart ist freischaffender Redaktor.

Unsere Expertin ► Sylvie Bonvin-Sansonnens ist Direktorin für Bildung und kulturelle Angelegenheiten und Präsidentin des Steuerungsausschusses. sylvie.bonvin-sansonnens@fr.ch

Unsere Expertin ► **Astrid Epiney** ist Rektorin der Universität Freiburg. astrid.epiney@unifr.ch

Unsere Expertin ➤ **Delphine Etienne- Tomasini** ist Rektorin der Pädagogischen Hochschule HEP | PH FR.
delphine.etienne-tomasini@edufr.ch

Unsere Expertin > Floriane Gasser ist-Vorsteherin des Amtes für Universitätsfragen und Präsidentin des Projektausschusses.

floriane.gasser@fr.ch

Das bleibt für die PH gleich

Für Studierende

Die Aufnahmebedingungen für angehende Lehrpersonen 1H bis 8H gelten weiterhin und der Abschluss bleibt ein Bachelor of Arts in Primary Education sowie ein Lehrdiplom für den Unterricht auf der gesamten Primarstufe. Der Anteil an Praktika und Berufspraxis sowie die Anforderungen an Praktikums-Lehrpersonen bleiben unverändert. «Heute und auch in Zukunft gehört die praktische Erfahrung zur DNA der Primarlehrer-Ausbildung», bekräftigt Delphine Etienne-Tomasini, Rektorin der PH. Zudem: Die Ausbildung von Primarlehrpersonen wird nicht zusätzlich akademisiert, bereits heute ist sie an einer tertiären Bildungseinrichtung angeboten.

Für Mitarbeitende

Das gesamte derzeitige Personal der PH erhält eine Gehaltsgarantie, das Projekt führt zu keinem Stellenabbau. «Die finanziellen Mittel der PH stehen in Zukunft der neuen Fakultät und der Universität zur Verfügung. Damit stärken wir die Positionierung der Uni als attraktive Ausbildungsstätte», sagt Floriane Gasser, Vorsteherin des Amtes für Universitätsfragen.

Die nächsten Schritte

2023 bis 2025

- ► Gesetzgebungsarbeiten, insb. Revision der Universitätsstatuten (2023/24)
- ► Erlass der Statuten der neuen Fakultät (2024)
- ► Anpassung der verschiedenen Reglemente (2024/25)
- ► Erarbeitung und Umsetzung der Detailkonzepte für den administrativen und akademischen Bereich

2025

- ► Gründung der neuen Fakultät am 1. August 2025
- Beginn des akademischen Jahres 2025/26 unter dem gleichen institutionellen Dach
- ▶ unifr.ch/go/regroupement



People & News

Sous la houlette du Président du Conseil national Martin Candinas, le Dies academicus s'est tenu le 15 novembre 2023. Au cours de cette cérémonie, chaque Faculté a remis le titre de docteur·e honoris causa à une personnalité marquante. La Faculté de droit a rendu hommage au parcours inhabituellement engagé du Professeur de droit français Pascal Ancel. La Faculté des lettres et des sciences humaines a choisi de distinguer Ann Blair, professeure d'histoire de l'Université de Harvard et experte de l'humanisme et de la culture savante des XVIe et XVIIe siècles. La Faculté des sciences économiques et sociales et du management a choisi d'honorer Mark Musen, professeur à l'Université de Standford, pour son exploration de l'interface entre l'intelligence artificielle et la santé humaine. La Faculté des sciences et de médecine a honoré Luca Vetterli, qui a consacré toute sa vie à la protection de la biodiversité et des zones humides en Suisse. La Faculté de théologie a, quant à elle, décidé de récompenser Hans-Dieter Mutschler, physicien, philosophe et théologien.

Cette journée a également été l'occasion de remettre de nombreux prix académiques. Le Prix Enseignement, offert par la Credit Suisse Foundation, est revenu à Dahlila Spagnuolo de la Faculté des lettres et des sciences humaines. Le Prix Leuba, décerné chaque année par l'Institut d'études œcuméniques à des travaux d'excellence promouvant l'encouragement d'une meilleure connaissance des différentes confessions chrétiennes, a été octroyé à Ciprian Sava-Popa pour son travail de master «Les limites de l'Eglise selon le concile Vatican II. Un regard orthodoxe sur un développement catholique». Institués en 1908, les Prix Vigener récompensent des travaux de diplôme, master ou doctorat de qualité exceptionnelle. A l'occasion du Dies academicus 2023, quatre facultés décernent des Prix Vigener: En droit, le prix récompense Sian Affolter pour sa thèse de doctorat intitulée «Der Umgang der Landwirtschaft mit der natürlichen Umwelt - de lege lata et de lege ferenda». En sciences économiques et sociales, le prix est décerné à Henrika Langen pour sa thèse de doctorat intitulée «Essays on Causal Inference», ainsi qu'à Romain Ducret pour sa thèse de doctorat «The Korean market and business groups: Three essays on the pricing, valuation and policies of Korean listed firms». En lettres et sciences humaines, la distinction revient à Serina Heinen et Antoine Sansonnens dont la thèse de doctorat est intitulée «Entretenir les possibles. Les praticiens de la relation d'aide face à la délicate mission de (rendre capable) des jeunes souffrant de troubles mentaux». En science et médecine, le prix est décerné à Timur Ashirov pour sa thèse intitulée «Functional Porous Materials and Membranes for Gas Capture and Separation Applications». Le Prix Chorafas, attribué par la fondation du même nom à de jeunes chercheuses et chercheurs du monde entier qui ont excellé par la qualité de leur travail dans les domaines de la biotechnologie, de la protection de l'environnement, des technologies d'information, des mathématiques, de la médecine, de la physique ou de la finance, a été remis cette année à Inès Arous pour sa thèse intitulée «Human-AI Collaborative Approaches for Open-Ended Data Curation». Le Prix du Prince du Liechtenstein distingue tous les deux ans l'excellence de travaux scientifiques qui se préoccupent d'une conception chrétienne du monde et de l'homme. Cette année, il a été attribué à Jean-François Corpataux de la Faculté des lettres et des sciences humaines pour son livre intitulé «L'image féconde. Art et dynastie à la Renaissance», et à Dominic Roser de la Faculté de droit pour son article «Effectiv

Altruism as Egyptian Gold for Christians» paru dans l'ouvrage Effective Altruism and Religion. Synergies, Tensions, Dialogue.

Neuria, la start-up fondée par l'équipe du Professeur Lucas Spierer, s'est vu attribuer un Future of Health Grant niveau 2 de 30'000 francs pour son travail sur les thérapies gamifiées pour améliorer les comportements individuels, tels que l'obésité, avec des interventions personnalisées adaptées aux besoins des patient·e·s. Le Docteur **Daniel Künzler** du Département Travail social, politique sociale et développement global, a été élu au Comité directeur du Comité de recherche RC 19 (Poverty, Social Welfare and Social Policy) de l'International Sociological Association. Le gouvernement du Québec a remis l'Ordre des francophones d'Amérique, qui a pour but de démontrer son attachement envers la promotion et la consolidation de l'espace francophone en Amérique, au Professeur Claude Hauser du Département d'histoire contemporaine.

L'Université de Fribourg se réjouit d'accueillir plusieurs nouvelles et nouveaux professeur·e·s: la Professeure ordinaire en tenure track **Géraldine Bugnon** au Département de travail social et politiques sociales; la Professeure assistante **Nesina Grütter** au Département d'études bibliques; le Professeur assistant **Alessandro Ianiro** à l'Institut Adolphe Merkle; la Professeure ordinaire **Sabine Krause** au Département des sciences de l'éducation; le Professeur assistant **Pascal Rey** au Département de droit privé; et le Professeur **Kilian Schindler** au Département d'anglais. Bienvenue à elles et eux.



Wolfgang Taube

Professor für Sportwissenschaften

Was langweilt Sie?

Ich habe mich schon lange nicht mehr gelangweilt. Es gibt immer etwas zu tun und wenn es mal Zeit zum Durchschnaufen gibt, habe ich gelernt, diese Zeit zu schätzen und zur Ruhe zu kommen

Was bereuen Sie in ihrem Leben?

Ich denke, dass ich das Wesentliche richtig gemacht habe. Manchmal hätte ich es vielleicht etwas «eleganter» machen können

Worin sollten Sie sich üben?

In kurzen Antworten;-)

Ihre wichtigste Charaktereigenschaft in Bezug auf Ihre Arbeit?

Zielstrebigkeit, Beharrlichkeit, Visionen und immer eine

gute Portion Humor

Welche Frage stellen Sie sich immer wieder?

Wann ist es endlich Zeit zum Mittagessen?

Was bringt Sie zum Weinen?

Es gibt genug Konflikte und Leid, wo einem die Tränen kommen können. Bei mir reicht aber manchmal schon ein schön emotionaler Film **Ei**

Wovon haben Sie keine Ahnung?

Von vielem. Aber ich versuche stetig dazuzulernen und ich kenne die Gefahr des Dunning-Kruger-Effektes gemäss dem Motto: «Das ist aber ganz schön viel MEINUNG für so wenig AHNUNG!»

Woran glauben Sie? An die Lernfähigkeit

Haben Sie einen Tick?Mir fällt jetzt zumindest keiner ein

Wovor haben Sie Angst? Dass meiner Familie etwas passiert

Möchten Sie lieber sterben oder als Tier weiterleben?

Lieber als Tier weiterleben. Zum Beispiel als Vogel mal das Fliegen erfahren. Sterben kann ich ja dann immer noch

Ihre liebste Tageszeit?

Zum Arbeiten oder Wandern der Morgen, für ein Mittagsschläfchen gleich nach dem Zmittag, für Sport der Nachmittag und für ein geselliges Abendessen der Abend

Eine Fähigkeit, die Sie gerne hätten?

Noch mehr «über den Dingen stehen»

Nous vous souhaitons une merveilleuse année 2024!

Wir wünschen Ihnen ein wunderbares Jahr 2024!

