



KI-Technologien in der Nähe von Datenquellen einzusetzen, kann Unternehmen und Personen Kontrolle über ihre Daten zurückgeben, erfordert aber den Aufbau komplexer Systeme. Diese Grand Challenge zeigt den – nicht unbedingt leichten – Weg zur individuellen Edge AI.

Künstliche Intelligenz (KI) bestimmt zunehmend unseren beruflichen und privaten Alltag. Dennoch sind KI und ihre bekannteste technische Umsetzung, das maschinelle Lernen (ML), stetiger Kritik ausgesetzt. Insbesondere wird kritisiert, dass das Training von ML-Modellen häufig auf zentralisierten Datenbeständen in der Cloud stattfindet. Dies impliziert, dass Nutzer*innen ihre Daten zunächst in die Cloud senden müssen und damit die Kontrolle über die eigenen Daten abgeben. Sowohl Unternehmen als auch Privatpersonen sind jedoch aus guten Gründen häufig nicht bereit, dies zu tun.

Da das Trainieren von ML-Modellen in vielen Fällen große Datenmengen benötigt, kann der notwendige Datentransfer zu weiteren Problemen führen. Dies ist insbesondere dann der Fall, wenn Ergebnisse in Echtzeit benötigt werden oder ein Unternehmen beziehungsweise eine Privatperson zwar Daten zur Verfügung stellen möchte, jedoch nicht über eine entsprechende Internetanbindung verfügt.

Aus diesen Gründen hat das Konzept der Edge-KI, also der Anwendung von KI-Technologien und insbesondere ML am Rande des Netzes, in letzter Zeit viel Aufmerksamkeit erlangt. Edge-KI basiert auf der Idee, dass neuartige KI-Technologien mittels Edge-Computing auf geeigneten lokalen Geräten in der Nähe von Datenquellen eingesetzt werden können, sodass Trainingsdaten nicht zentral, zum Beispiel in der Cloud, gesammelt werden müssen. Dies führt zu einem dezentraleren Ansatz bei der Anwendung von KI.

Grundlegende Konzepte der Edge-KI werden heute bereits eingesetzt. Dennoch müssen zunächst noch eine Vielzahl von Forschungsfragestellungen beantwortet werden, bevor die Technologie breite Anwendung finden kann. Diese reichen von speziell angepassten ML-Algorithmen für leichtgewichtige IT-Umgebungen über nachvollziehbare KI bis hin zur Benutzerfreundlichkeit.

Das ist die Challenge

Trotz des Einflusses von Künstlicher Intelligenz (KI) und maschinellem Lernen (ML), ist die Kritik vielfältig und umfasst technische, gesellschaftliche und staatliche Aspekte.

Ein wesentlicher Kritikpunkt an ML ist die Tendenz zu natürlichen Monopolen. In vielen Fällen werden für die Anwendung von ML sehr große Datenmengen für das Training sowie sehr große Mengen an Rechenleistung benötigt. Dies führt natürlich dazu, dass sich ML-Expertise in Unternehmen ansammelt, die bereits über große Datenmengen und sehr große Rechenzentren verfügen. Es überrascht daher nicht, dass Unternehmen wie Google oder Facebook ebenfalls stark im Bereich der KI engagiert sind. Selbst neuere Akteure wie OpenAI werden von Unternehmen wie Microsoft und Amazon unterstützt und profitieren von deren Daten und Rechenleistung.

Viele Unternehmen und Menschen sind jedoch nicht sehr erpicht darauf, ihre Daten an solche Unternehmen zu senden, da sie dadurch die Kontrolle über ihre Daten verlieren. Daher können zentralisierte ML-Ansätze zu einer Bedrohung für die Datensouveränität werden.

Aus technischer Sicht kann das Hochladen sehr großer Datenmengen an eine zentrale Stelle für die Zwecke von ML problematisch werden. Selbst in einem Land mit relativ moderner Infrastruktur wie Deutschland haben nicht alle Unternehmen Zugang zu einer Hochgeschwindigkeits-Internetverbindung. Dies gilt insbesondere für ländliche Gebiete und für kleine und mittlere Unternehmen, die das Rückgrat der deutschen Wirtschaft bilden. Wenn die Internetverbindung zu einem Engpass wird, könnte dies die Unternehmen davon abhalten, KI-Funktionalitäten zu nutzen. Darüber hinaus erfordern verschiedene KI-Anwendungsfälle wie autonome Fahrzeuge oder die industrielle Fertigung, dass die Verarbeitung und Entscheidungsfindung in Echtzeit erfolgt. Das könnte problematisch werden, wenn das Hoch- und Runterladen von Daten in die Cloud mit zu hohen Latenzzeiten verbunden ist.

Edge AI führt zu einem dezentraleren Ansatz bei der Anwendung von KI, indem das Training und die Verwendung von Modellen am Rande des Netzwerks geschehen. Eine andere Richtung der Edge AI besteht darin, ML-Modelle weiterhin in der Cloud zu trainieren, aber die Modelle in der Nähe der Datenquellen anzuwenden. Dabei ist jedoch zu berücksichtigen, dass die Vorteile in puncto Datenschutz während der Trainingsphase verpuffen.

Während das Grundkonzept der Edge AI vielversprechend ist, führt es auch zu komplexen neuen Forschungsfragen. Diese ähneln teilweise den Fragen, die auch in der zentralisierten KI und im generischen Edge-Computing gelöst werden müssen, sind aber teilweise Edge AI-spezifisch.

Wie aus vielen anderen Bereichen der Informatik bekannt ist, kann die Dezentralisierung und Verteilung von Aufgaben sehr vorteilhaft sein, führt aber auch zu einer höheren Systemkomplexität. Dementsprechend erbt Edge AI die Herausforderungen der zentralisierten KI und führt durch die steigende Systemkomplexität neue ein. So geht eine höhere Komplexität in der Regel mit einer größeren Bedrohungsfläche einher und führt zu zusätzlichen Sicherheitsherausforderungen, nämlich der Gewährleistung der Vertraulichkeit, Integrität und Verfügbarkeit von Daten und des Edge AI-Systems selbst. Edge AI ist, insbesondere wenn sie mit dezentralem Lernen gekoppelt ist, anfälliger für die Einbindung von Hintertüren als zentralisierte KI. Da Edge AI-Systeme kritische Komponenten in künftigen Infrastrukturen und vernetzten Diensten sein könnten, müssen sie äußerst widerstandsfähig gegen Ausfälle und Angriffe sein.

Ziel muss es sein, Edge AI-Technologien zu etablieren, die leichtgewichtig im Hinblick auf den Ressourcen- und Energieverbrauch, aber dennoch leicht kontrollierbar, sicher und widerstandsfähig genug sind, um in einem breiten Spektrum von Anwendungsfällen und auch von Nutzenden eingesetzt zu werden, die keine KI-Experten sind.

Darum ist es wichtig, das Problem anzugehen

Um die Komplexität und die Schwierigkeiten bei der Verwirklichung von Edge AI zu verdeutlichen, lohnt sich der genauere Blick auf einige der wichtigsten Teilherausforderungen. Sie zeigen, dass viele Teildisziplinen der Informatik sowie andere Disziplinen ihre Kräfte bündeln müssen, um die großen Herausforderungen von Edge AI zu lösen.

Erstens gibt es in der ML einen Trend zu immer komplexeren Modellen. Bereits relativ kleine neuronale Netzmodelle können eine sechsstellige Anzahl von Parametern haben, und sehr große Modelle können Milliarden von Parametern umfassen. Dementsprechend gibt es auch einen Trend zu immer höheren Anforderungen an die Rechenleistung. Dies kollidiert mit der in der Regel geringeren Rechenleistung der am Rande des Netzes verfügbaren Geräte. Ein Lösungsansatz wäre, weniger komplexe Modelle zu trainieren. Wenn allerdings weniger Parameter in einem ML-Modell berücksichtigt werden können, kann dies zu geringerer Genauigkeit und sogar zu Unbrauchbarkeit führen. Auch wenn einige Kompromisse zwischen Genauigkeit und Ressourcenverbrauch akzeptabel sein mögen, ist es nicht trivial, den goldenen Mittelweg zwischen diesen beiden widersprüchlichen Anforderungen zu finden. Daher sind Forschungsanstrengungen erforderlich, um den Ressourcenverbrauch klassischer ML-Probleme zu optimieren und gleichzeitig eine ausreichende Genauigkeit zu erzielen.

Zweitens können Edge AI-Systemlandschaften sehr komplex werden, vor allem wenn ein dezentraler Ansatz für ML wie Federated Learning (FL) angewandt wird. Bei FL werden kritische Rohdaten nicht an eine zentrale Einheit gesendet, sondern nur ein erlerntes lokales Modell wird mit dem Aggregator geteilt. Dies führt zwar zu Vorteilen hinsichtlich des Kommunikationsaufwands und des Datenschutzes, aber die resultierenden Systeme können Tausende von teilnehmenden Clients enthalten, die alle ihre eigenen lokalen Modelle liefern, und Millionen von Datenquellen. Um zu vermeiden, dass dies zu einem Overhead führt, der die Vorteile von Edge AI zunichtemacht, ist es notwendig, neue Ansätze zu finden, die helfen, FL-Systemlandschaften im Hinblick auf verschiedene Ziele zu optimieren, zum Beispiel Latenz, Energieverbrauch oder Genauigkeit. Teilweise werden die entsprechenden Fragestellungen direkt aus dem allgemeineren Bereich des Edge-Computings übernommen, wo der Aufbau von effizienten Systemlandschaften ebenfalls von großer Bedeutung ist. Edge AI bringt jedoch spezifische Kommunikationsmuster und Anforderungen hinsichtlich der Einrichtung von Knotenkohorten mit sich, die über generische Edge-Computing-Systemlandschaften hinausgehen.

Drittens müssen die Mobilfunknetze der nächsten Generation berücksichtigt werden, da die mobile Nutzung von Edge AI erfordert, dass der Netzwerkrand zum Konvergenzpunkt für die Datenverarbeitung wird. Network Functions Virtualization (NFV) muss mit grundlegenden Microservice-Prinzipien kombiniert werden, um sicherzustellen, dass einzelne Edge AI-Dienste bis hin zu kompletten Edge AI-Systemlandschaften sowohl leistungsstark als auch widerstandsfähig sind. Um Edge AI auf breiter Ebene für die mobile Nutzung zu realisieren, ist es notwendig, die Idee der (5G-)Edge-Server zu erweitern, da diese heute nur in begrenztem Umfang verfügbar sind. Es ist notwendig, die Möglichkeit zu schaffen, jede (mobile) Edge-Ressource zu virtualisieren. Dies erfordert Virtualisierungsmechanismen, die über die heute gängigen Microservice-Technologien hinausgehen, da diese für den Einsatz auf vielen Edge-Geräten zu schwergewichtig sind.

Viertens verlangen viele Nutzende von ML eine erklärbare KI: Nutzende wollen verstehen, wie die Ergebnisse zustande gekommen sind, anstatt auf den Black-Box-Ansatz zu vertrauen, den ML normalerweise verfolgt. Bei der Anwendung von FL wird beispielsweise eine weitere Abstraktionsebene bei der Erstellung eines globalen ML-Modells hinzugefügt, da lokale Modelle aggregiert werden müssen. Dies macht es noch komplexer, erklärbare KI zu gewährleisten, da gängige Ansätze wie die schichtweise Relevanzübertragung oder die kontrafaktische Methode erweitert werden müssen, um die lokalen und globalen Modelle und die Tatsache zu berücksichtigen, dass möglicherweise nur ein Teil der Rohdaten verfügbar ist. Erweiterte und neuartige Ansätze für erklärbare KI werden auch bei der Optimierung von ML-Systemen und Trainingsverfahren hilfreich sein.

Fünftens führt Edge AI zu neuen Forschungsfragen im Bereich der Sicherheit und Belastbarkeit. Der Datenschutz könnte durch die Lokalisierung von Daten verbessert werden. Allerdings treten auch neue Angriffsvektoren auf. So kann ein Angreifer beispielsweise die Qualität eines Modells durch Data Poisoning beeinträchtigen. Bei zentralisiertem ML kann Data Poisoning leichter erkannt werden, zum Beispiel wenn die Daten eindeutig nicht dem entsprechen, was realistischerweise erwartet werden kann. Bei Edge AI ist die lokale Datenmenge möglicherweise zu gering, um dies zu erkennen. Bei dezentralem ML tauchen neue Angriffsvektoren wie Model Poisoning auf, bei dem Clients falsche lokale Modelle an einen Aggregator liefern. Dies zeigt, dass neuartige Sicherheits- und Ausfallsicherheitsmechanismen erforderlich sind, die Angriffe frühzeitig erkennen können, sodass schädliche Folgen minimiert werden können.

Nicht zuletzt sollte Edge AI leicht anwendbar sein. Selbst Nutzende mit wenig ML-Kenntnissen sollten in der Lage sein, Datenquellen zu integrieren und Modelle zu trainieren und anzuwenden. Im Idealfall ist die Nutzungserfahrung ähnlich wie bei modernen ML-Anwendungen wie ChatGPT oder Stable Diffusion. Es werden neuartige Ansätze benötigt, die es ermöglichen, Edge AI-Systeme nach dem Plug-and-Play-Prinzip einzurichten und so die Planungs- und Optimierungsaufgaben für Systementwickelnde und -betreibende zu minimieren.

In diesen Zeithorizont ist eine Lösung zu erwarten

Zwar gibt es bereits erste Lösungen für Edge AI-Probleme, doch stehen sowohl die Forschung als auch die Übernahme durch die Industrie noch ganz am Anfang. Trotz der raschen Verbreitung von ML in Gesellschaft und Industrie werden noch mindestens fünf Jahre vergehen, bis Edge AI auf breiter Front eingesetzt werden kann. Sicherlich werden nicht alle Forschungsfragen in diesem Bereich bis 2028 oder 2029 beantwortet sein. Edge AI wird in den nächsten fünf Jahren dennoch einige große Sprünge machen.

Diese Anstrengungen wurden bereits in die Lösung investiert

Edge AI ist ein sehr junges Thema, das jedoch mit der Verfügbarkeit leistungsfähiger und mobiler Edge-Geräte in den vergangenen zehn Jahren sowie der zunehmenden Menge an verfügbaren Daten, insbesondere im Internet of Things, stark an Dynamik gewonnen hat. Grundlegende Ergebnisse haben bereits zu anwendbaren Lösungen geführt, zum Beispiel in den Bereichen Überwachung und Sicherheit oder Industrie 4.0, und in Start-ups, die explizit auf die Bereitstellung von Edge AI-Lösungen abzielen. Dennoch gibt es weiteren Forschungs-

und Technologieentwicklungsbedarf, vor allem in Bezug auf die Effizienz von KI-Algorithmen für den Einsatz auf Edge-Geräten.

Welche Disziplinen an dieser Challenge zusammenarbeiten müssen

Die Vielzahl möglicher Forschungsfragen spiegelt sich in den vielen wissenschaftlichen Konferenzen und Fachzeitschriften wider, die heutzutage zu Beiträgen im Bereich Edge AI aufrufen. Dazu gehören Spitzenkonferenzen in den Bereichen Maschinelles Lernen, verteilte und vernetzte Systeme, Sicherheit und Data Mining, die nach grundlegenden Forschungsbeiträgen zu verschiedenen Teilproblemen der Edge AI suchen. Darüber hinaus ist Edge AI auch ein Thema in Feldern, in denen sie hauptsächlich angewendet wird, zum Beispiel in der Computer Vision oder der Robotik.

Dies zeigt, dass die Anstrengungen verschiedener Teildisziplinen der Informatik notwendig sind, um Edge AI weiter voranzutreiben. Für einige der Forschungsfragen ist Fachwissen aus anderen Forschungsdisziplinen erforderlich, etwa aus der Mathematik für die Grundlagen von ML-Algorithmen oder aus der Elektrotechnik für spezielle Edge AI-Hardware. Außerdem ist für die Lösung bestimmter Forschungsfragen Domänenwissen erforderlich, beispielsweise für die Anwendung von Edge AI in der Fertigung. Dies zeigt, dass die große Herausforderung der Edge AI nur auf interdisziplinäre Weise gelöst werden kann – in der Informatik und darüber hinaus.

So könnte das Ziel aussehen

Die Grand Challenge Edge AI ist gelöst, wenn es möglich ist, KI am Rande des Netzes auf nahtlose und einfach zu bedienende Weise anzuwenden. Aufgrund der verschiedenen Disziplinen und Aspekte, die bei Edge AI eine Rolle spielen, gibt es kein einzelnes Kriterium, mit dem sich zeigen ließe, dass die Grand Challenge gelöst ist.

Ein Beispielszenario: Unternehmen, die in der Fertigung tätig sind, haben vielleicht nur eine oder zwei Maschinen eines bestimmten Typs in ihrer Werkstatt stehen. Um genaue ML-Modelle, zum Beispiel für die Qualitätskontrolle, zu erstellen, reicht die Menge der von diesen Maschinen stammenden Daten möglicherweise nicht aus. Dies stellt ein Unternehmen vor die Wahl, einerseits die Vorteile von ML nicht zu nutzen, andererseits mit weniger genauen Modellen zu arbeiten oder aber die Daten mit Dritten, also sehr oft mit dem Hersteller der Maschinen, zu teilen, um ein globales Modell zu erstellen. Während die dritte Option zu einem genauen Modell führen kann, wird die Weitergabe von Daten an Dritte oft skeptisch gesehen, da die Daten wichtige Einblicke in Geschäftsgeheimnisse geben können, zum Beispiel über Betriebszeiten und die Auslastung einer Maschine. Dies hängt mit dem zuvor erwähnten Kontrollverlust bei der Weitergabe von Daten an die Cloud zusammen. Auch können Unternehmen nicht unbedingt sicher sein, für welche anderen Zwecke ihre Daten verwendet werden.

Als Alternative könnten Unternehmen Edge AI-Methoden wie beispielsweise FL nutzen. Dazu muss eine Nutzende mit Domänenwissen über ein Toolset verfügen, um heterogene Datenquellen einfach in den Trainingsprozess zu integrieren, sodass die Daten automatisch in ein Format umgewandelt werden, das zum Trainieren eines ML-Modells und später beispielsweise für die vorausschauende Wartung verwendet werden kann. Basierend auf

ihren Entscheidungen wird automatisch eine ML-Pipeline generiert, die zur Bereitstellung und Optimierung der Edge AI-Systemlandschaft dient. Trotz der großen Datenmengen, die in diesem Szenario verwendet werden, kann das Training auf relativ simplen Geräten, zum Beispiel Einplatinencomputern wie Raspberry Pis, durchgeführt werden. Dies ist möglich, weil die angewandten ML-Algorithmen für die Ausführung auf solchen Geräten optimiert sind und leichtgewichtige Mechanismen zur Virtualisierung aller Edge-Geräte sowie der Kommunikation vorhanden sind. Darüber hinaus kann das Training – selbst für einen einzelnen Client – auf verschiedene Geräte entlang des Cloud-to-Edge-Kontinuums verteilt werden, wiederum ohne manuelle Eingriffe der Nutzenden oder Betreibenden, sondern ausschließlich auf der Grundlage benutzerdefinierter Anforderungen wie hohe Genauigkeit oder sehr geringe Latenz.

Nach der Erstellung des lokalen FL-Modells wird das Modell in einen zentralen Aggregator hochgeladen, der das globale Modell erstellt und es an die Kunden zurückschickt. Vielleicht führt das globale Modell dennoch für ein Unternehmen nicht zu guten Ergebnissen. Die Gründe dafür werden der Nutzenden jedoch auf der Grundlage neuartiger Techniken nachvollziehbar gemacht, sodass sie das globale Modell so optimieren kann, dass es zu einer höheren Genauigkeit für das Unternehmen führt. Außerdem werden die Erklärungen an den Aggregator weitergeleitet, der diese Daten nutzen kann, um Kohorten von Unternehmen zu identifizieren, die ihr eigenes gemeinsames „globales“ Modell erhalten sollen, wodurch das Problem der nicht unabhängigen und identisch verteilten (non-IID) Daten gelöst wird.

All diese Funktionalitäten sind sicher und belastbar, da Sicherheit und Datenschutz von Anfang an, also während des gesamten Lebenszyklus der Edge AI-Software und der Systemlandschaften, bei der Konzeption berücksichtigt werden.

Sobald ein solches Szenario realistisch anwendbar ist, kann die große Herausforderung der Edge AI als „gelöst“ bezeichnet werden, auch wenn natürlich weitere Entwicklungen erforderlich sind.

Welche sozialen, gesellschaftlichen oder ökonomischen Probleme sich mit der Grand Challenge adressieren lassen

Wie zuvor beschrieben, trägt Edge AI zur Lösung von zwei großen Problemen bei, die derzeit in den Medien, der Politik und der Industrie breit diskutiert werden. Sie ermöglicht es, KI zu nutzen und damit von ihren vielen möglichen Anwendungsbereichen zu profitieren, ohne Rohdaten mit Dritten zu teilen.

Vor allem in Deutschland werden Datenschutz und Datensouveränität oft als große Hürden für die Nutzung und Entwicklung neuer KI-Technologien gesehen. Während größere Unternehmen über die notwendigen Mittel verfügen, um eine Abwägung zwischen Datensouveränität und neuartigen Funktionalitäten vorzunehmen, ist dies vor allem bei kleinen und mittleren Unternehmen oft nicht der Fall. Dies trifft natürlich auch auf die meisten Privatpersonen zu. Daher ist die Bereitstellung von sofort einsetzbaren Lösungen für Edge AI auf der Grundlage von Forschung und technologischer Entwicklung ein sehr wichtiges Mittel, um die breite Einführung von ML zu gewährleisten.

Aus wirtschaftlicher Sicht schätzt Statista, dass der weltweite KI-Markt jährlich von 208 Milliarden USD im Jahr 2023 auf 1.850 Milliarden USD im Jahr 2030 wachsen wird, was einer durchschnittlichen jährlichen Wachstumsrate von 32,9 Prozent entspricht. Wenn deutsche Unternehmen auf den Einsatz dieser Technologien verzichten, besteht ein großes Risiko, dass sie Marktanteile verlieren.

Obwohl hier ein industrieller Anwendungsfall beschrieben wurde, ist Edge AI auch in anderen Bereichen mit sehr großen gesellschaftlichen Auswirkungen wichtig. So nimmt beispielsweise der Anteil älterer Menschen, die allein leben, stetig zu. Edge AI kann eingesetzt werden, um ältere Menschen im Blick zu behalten, zum Beispiel um mithilfe einer intelligenten Kamera zu erkennen, ob eine Person gestürzt ist und hilflos ist. Viele Menschen sind vielleicht nicht damit einverstanden, Videomaterial von zu Hause auf einen öffentlichen Cloud-Anbieter hochzuladen, um eine Notsituation zu erkennen. Mit Edge AI ist es möglich, die notwendige Videoverarbeitung und -identifizierung vor Ort auf einem Edge-Gerät durchzuführen, so dass das Videomaterial das Haus der gefilmten Person niemals verlässt. Auch dies ist nur ein Beispiel für die vielen potenziellen Vorteile, die Edge AI mit sich bringen kann.

Über die Autor*innen

Prof. Dr. Christian Becker leitet das Institut für Parallele und Verteilte Systeme an der Universität Stuttgart. Seine Forschung beschäftigt sich mit unterschiedlichen Fragestellungen aus dem Bereich der verteilten Systeme, inklusive adaptiver Edge AI-Landschaften.

Prof. Dr. Mathias Fischer leitet die Arbeitsgruppe Rechnernetze an der Universität Hamburg. Der Fokus seiner Forschung liegt auf Sicherheit und Resilienz in vernetzten und verteilten Systemen.

Prof. Dr.-Ing. Stefan Schulte forscht zu Optimierung und Interoperabilität in verteilten Systemen. Er leitet das Institut für Data Engineering an der Technischen Universität Hamburg und ist Sprecher der GI-Fachgruppe Kommunikation und Verteilte Systeme.

Prof. Dr.-Ing. Lars Wolf ist Leiter des Instituts für Betriebssysteme und Rechnerverbund, Abteilung Connected and Mobile Systems, an der Technischen Universität Braunschweig. Ein besonderer Schwerpunkt seiner Forschung liegt auf der Verwendung von Edge AI in drahtlosen Sensornetzen.

Prof. Dr.-Ing. Falko Dressler ist Inhaber des Lehrstuhls für Telekommunikationsnetze an der Technischen Universität Berlin. Er arbeitet an unterschiedlichen Forschungsfragen im Bereich mobiler Kommunikationsnetze, inklusive der Nutzung von KI in solchen Netzen.

Unterstützende GI-Gliederungen und Mitautor*innen

Fachgruppe Kommunikation und Verteilte Systeme (KuVS)

Fachgruppe Erkennung und Beherrschung von Vorfällen der Informationssicherheit (SIDAR)

Weiterführende Literatur

Ding, A. Y., Peltonen, E., Meuser, T., Aral, A., Becker, C., Dustdar, S., Hiessl, T., Kranzlmüller, D., Liyanage, M., Maghsudi, S., Mohan, N., Ott, J., Rellermeyer, J. S., Schulte, S., Schulzrinne, H., Solmaz, G., Tarkoma, S., Varghese, B., Wolf, L. (2022). Roadmap for edge AI: a Dagstuhl perspective. *SIGCOMM Computer Communication Review*, 52, 1, 28–33.
<https://doi.org/10.1145/3523230.3523235>