Analyse potenzieller Cybersicherheitsrisiken beim Einsatz von KI im Zukunftsmodell des Schweizer Stromökosystems

Autoren: Dr. David M. Sommer, Felix Huber, Pascal Schöni, Dr. Raphael M. Reischuk

Version 1.0, 25. Februar 2025

Inhaltsverzeichnis

Nationales Testinstitut für Cybersicherheit NTC

Das Nationale Testinstitut für Cybersicherheit NTC trägt zur Sicherheit und zur digitalen Souveränität der Schweiz bei, indem es kritische Schwachstellen vorausschauend erkennt und deren Behebung fördert. Als nicht gewinnorientierter Verein mit Sitz in Zug folgt das NTC den Prinzipien der Unabhängigkeit und Objektivität. Es führt Cybersicherheitsprüfungen von vernetzten Infrastrukturen, Geräten und Anwendungen durch, die für Gesellschaft und Wirtschaft von grosser Bedeutung sind.

Kontakt:
Andreas W. Kaelin
office@ntc.swiss

Danksagung

Wir danken unseren Expertinnen und Experten aus den Bereichen Cybersicherheit, Digitalisierung und Künstliche Intelligenz KI für die wertvollen Beiträge zu dieser Publikation.

Gerhard Andrey
Nationalrat GRÜNE und Unternehmer bei Liip

Lukas Mäder
Redaktor Technologie bei der NZZ

Dr. Serge Droz
Senior Technical Advisor beim EDA

Markus Rööslı
Direktor Polizeitechnik und -informatik Schweiz (PTI)

Adrienne Fichter
Tech-Journalistin bei der Republik

David Rosenthal
Team Head und Partner bei VISCHER

Dr. Christian Folini
Cybersicherheitsexperte

Martin Steiger
Rechtsanwalt für Recht im digitalen Raum bei der Steiger Legal AG

André Golliez
Präsident der Swiss Data Alliance

Dr. Florian Tramèr
Assistant Professor of Computer Science an der ETH Zürich

Dr. Vincent Lenders
Leiter des Cyber-Defence Campus bei armasuisse

Livia Walpen
Senior Policy Advisor International Relations beim BAKOM

1 Management Summary

Der zu erwartende Anstieg des Einsatzes von Künstlicher Intelligenz (KI) in kritischen Infrastrukturen birgt neue Herausforderungen für die Cybersicherheit und digitale Souveränität der Schweiz.

Das Nationale Testinstitut für Cybersicherheit NTC hat deshalb eine Risikobewertung für den KI-gestützten Ausbau des Schweizer Stromökosystems durchgeführt, um potenzielle Bedrohungen frühzeitig zu erkennen und Empfehlungen abzuleiten.

Im Rahmen der Analyse wurden die folgenden acht Risiken für ein mögliches Zukunftsmodell des Schweizer Stromökosystems identifiziert und mit Cyber-, Digitalisierungs-, und KI-Expertinnen und -Experten diskutiert.

1. **Erhöhte Angriffsfläche kommerzieller Endkundenprodukte**
2. **Kaskadeneffekte aufgrund hoher Komplexität**
3. **Unabsichtliche KI-Fehler**
4. **Herstellerabhängigkeiten**
5. **Abfluss und Missbrauch sensitiver Daten**
6. **Störungen bei der Datenübertragung**
7. **Backdoors und gezielte Manipulation**
8. **Absichtliche Manipulation der Datengrundlage**

Die Expertinnen und Experten haben in den Gesprächen deutlich die Wichtigkeit einer Analyse des sich verändernden Stromökosystems hervorgehoben – sei dies aufgrund von zunehmender Digitalisierung, Vernetzung oder dem Einsatz von KI. Unter den Expertinnen und Experten bestand Konsens über die Relevanz und Dringlichkeit der identifizierten acht Risiken. Sie haben bekräftigt, dass es sich um ernstzunehmende Risiken handelt, die berücksichtigt werden müssen. Entsprechend wurden viele Risiken als „hoch" oder „sehr hoch" eingestuft. Die Ergebnisse unterstreichen die Notwendigkeit, dass das NTC die durch die KI hervorgerufenen Herausforderungen intensiv und proaktiv behandelt.

Am höchsten bewertet haben Expertinnen und Experten das Risiko von kommerziellen Endkundenprodukten, die in einer sehr grossen Zahl in das Kommunikationsnetz des Stromsystems eingebunden sind. Kommerzielle Endkundenprodukte sind Produkte, die für den direkten Verkauf an Verbrauchende (Endkundinnen und -kunden) bestimmt sind. Diese Produkte werden für den persönlichen oder häuslichen Gebrauch entwickelt und vertrieben, im Gegensatz zu Produkten, die für Unternehmen oder den industriellen Gebrauch bestimmt sind. Kommerzielle Endkundenprodukte umfassen Geräte, wie intelligente Haushaltsgeräte, die mit dem Internet verbunden sind. Dies können Waschmaschinen, Trockner oder Wechselrichter von Solarstromanlagen sein. Schwachstellen aufgrund unzureichender Cybersicherheit und fehlenden Sicherheitsupdates könnten ausgenutzt werden, um das Stromnetz zu destabilisieren, falsche Informationen einzuspeisen oder destruktive Verbrauchs- und Produktionsmuster zu erzeugen.

Dieses Risiko wird durch den Einsatz von KI erhöht. Einerseits besteht eine Abhängigkeit zu den wenigen Anbietern von KI-Modellen, welche absichtlich oder unabsichtlich destruktive Verbrauchs- oder Produktionsmuster erzeugen könnten. Andererseits steigt die Komplexität, wodurch es für Menschen zunehmend schwierig wird, das Zusammenspiel verschiedener, miteinander vernetzter Komponenten vollständig nachzuvollziehen, um Vorfälle antizipieren und abwenden zu können.

11 von 12 Expertinnen und Experten haben sich bei der Risikobewertung für eine der beiden höchsten Wahrscheinlichkeiten («wahrscheinlich» bis «sehr wahrscheinlich») ausgesprochen und die Auswirkungen als «erheblich» bis «kritisch» eingestuft. Die Expertinnen und Experten stimmten zu, dass diese Geräte aufgrund des Zeitdrucks für eine schnelle Markteinführung oft nicht ausreichend auf ihre Cybersicherheit getestet werden.

Empfehlungen

Den Cybersicherheitsrisiken, die durch die KI hervorgerufen werden, muss mit der notwendigen Sensibilisierung und Regulierung, den notwendigen Mitteln, sowie wiederholender Cybersicherheitsprüfungen begegnet werden:

- Koordinierte und wiederholte proaktive Cybersicherheitsprüfungen werden empfohlen – insbesondere bei kommerziellen Endkundenprodukten.
- Hinsichtlich der zunehmenden Komplexität von KI-Systemen sind regelmässige Schulungen der Anwendenden und Betreibenden sowie Sensibilisierungsmassnahmen unerlässlich.
- Digitale Infrastrukturen mit grosser Bedeutung für die Wirtschaft und Gesellschaft müssen mit den Mitteln ausgestattet werden, um bestehenden und durch KI zusätzlich hervorgerufenen Cybersicherheitsrisiken entgegenzutreten.
- Durch eine diversifizierte Beschaffungsstrategie können zudem Abhängigkeiten von den wenigen KI-Modell-Anbietern und geopolitische Risiken reduziert werden. Das Ausfallrisiko lässt sich zudem im Rahmen einer geplanten Diversifizierung senken. Open Source Software bieten mögliche Alternativen.

2 Einführung und Motivation

Durch die rasant zunehmende autonome Aufgabenerfüllung durch die Künstliche Intelligenz (KI) entstehen neue Risiken für die Sicherheit und digitale Souveränität der Schweiz. Das Nationale Testinstitut für Cybersicherheit NTC will diese frühzeitig erkennen, analysieren und antizipieren.

KI - hier definiert als Technologie zur automatisierten und teil-autonomen Ausführung von Aufgaben - bietet transformative Chancen zur Optimierung komplexer Infrastrukturen. Gleichzeitig birgt sie erhebliche Risiken, da es durch die zunehmende Automatisierung und Vernetzung für den Menschen schwieriger wird, die Funktionsweise dieser Systeme und die Vorhersehbarkeit ihrer Auswirkungen zu verstehen. Der vorliegende Bericht untersucht die Sicherheitsrisiken, die mit dem Einsatz von KI verbunden sind, anhand eines Zukunftsszenarios: dem KI-unterstützten, intelligenten Schweizer Stromökosystem. Er gibt Empfehlungen zur Risikominderung und zur sicheren Gestaltung kritischer Infrastrukturen. Der Bericht soll als Grundlage für eine weiterführende Diskussion über den sicheren Einsatz von KI in und für kritische Infrastruktur dienen.

3 Vorgehen

Das NTC führte eine Risikobeurteilung zum Einsatz von KI im Schweizer Stromökosystem durch, um potenziell systemdestabilisierende Effekte zu identifizieren. Die Analyse stützte sich auf Recherchen und bestehende Studien [1], welche einerseits darauf deuteten, dass in der Schweiz erfreulicherweise vorsichtig mit dem Einsatz von KI in kritischen Bereichen umgegangen wird. Gleichzeitig bestätigten sie das NTC in den Bedenken hinsichtlich möglicher Risiken beim Einsatz von KI. Um die systemischen Risiken für Wirtschaft und Gesellschaft, welche durch KI entstehen bzw. erhärtet oder abgeschwächt werden, zu analysieren, hat das NTC das Schweizer Stromökosystem untersucht.

Das Szenario des Schweizer Stromökosystems wurde gewählt, da es repräsentativ, zugänglich und häufig intuitiv verständlich ist. Zudem ist die Eintrittswahrscheinlichkeit von KI-induzierten Sicherheitsrisiken als hoch einzuschätzen. Das Beispiel des beinahe-Blackouts im April 2024, ausgelöst durch falsche Produktionsprognosen (Blick, 2024) unterstützt diese Einschätzung. Es wird erwartet, dass die anhand dieses Szenarios identifizierten Risiken auch auf andere Anwendungsbereiche von KI übertragen werden können.

Die Risikobewertung wurde nach ISO 31000 [2] und das Threat-Modelling nach dem STRIDE-Modell [3] durchgeführt. Anhand einer einfachen Risikomatrix wurden potenzielle Risiken basierend auf deren Eintrittswahrscheinlichkeit [4] und Auswirkungen [5] von Cyber-, Digitalisierungs-, und KI-Expertinnen und -Experten eingeschätzt.

4 Szenario

Die Entwicklung des Schweizer Stromökosystems hin zu **Intelligenten Netzen** (Smart Grids) ist ein zentraler Bestandteil der Energiestrategie 2050 und spielt eine Schlüsselrolle bei der Transformation des Schweizer Energiesystems (BFE, 2021).

Gemäss der Studie vom VSE (2022) zeigt ein Blick in das Jahr 2050, dass der Strombedarf in der Schweiz zunehmen wird und wir unter anderem ohne massive Steigerung der Energieeffizienz die Energie- und Klimaziele nicht erreichen werden (VSE, 2022). Der steigende Anteil an dezentraler Stromerzeugung und die Notwendigkeit, die gesamte Energieeffizienz in der Schweiz zu erhöhen, führen zu zahlreichen neuen Herausforderungen für die Stromnetze (BFE, 2021).

Intelligente Netze – Smart Grids – tragen dazu bei, diesen Herausforderungen zu begegnen (ibid.). Ein Smart Grid ist ein intelligentes Stromnetz, das digitale Informations- und Kommunikationstechnologien nutzt, um den Stromtransport effizient zu überwachen und zu steuern. Es koordiniert Erzeugende, Netzbetreibende, Endverbrauchende und Akteure des Strommarktes, um Kosten und Umweltauswirkungen zu minimieren sowie Zuverlässigkeit, Flexibilität und Stabilität zu maximieren (International Energy Agency, n.d.).

Die Transformation hin zu Smart Grids wird durch technologische Innovationen und gesetzliche Rahmenbedingungen, wie das Bundesgesetz über eine sichere Stromversorgung mit erneuerbaren Energien (Mantelerlass), unterstützt (BFE, 2024). Mit diesem Gesetz ist es zukünftig möglich, die lokal selbst erzeugte Elektrizität (z.B. durch eine Solaranlage) innerhalb eines Quartiers oder einer Gemeinde über das öffentliche Netz zu vermarkten (UVEK, 2024). Diese Gemeinschaften fördern die Nutzung von (Dach-)flächen für Photovoltaikanlagen und bieten eine Plattform für so genannte «Prosumer» (gleichzeitig Konsumierende und Produzierende), Speicherbetreibende, Endverbrauchende und Erzeugende. Voraussetzungen für die Teilnahme in lokalen

[1] AlgorithmWatch / CH bildet eine zuverlässige Quelle, welche sich mit den Einsatzorten von KI-Systemen beschäftigt: https://algorithmwatch.ch/de/atlas-der-automatisierung/

[2] ISO 31000 für Risk Management: https://www.iso.org/iso-31000-risk-management.html

[3] STRIDE ist ein Framework zur Identifizierung und Klassifikation von Sicherheitsbedrohungen in IT-Systemen. https://learn.microsoft.com/en-us/previous-versions/commerce-server/ee823878(v=cs.20)

[4] Die Wahrscheinlichkeiten wurden anhand der folgenden Skala bewertet: 5: Sehr wahrscheinlich (91% oder mehr Chance des Eintretens), 4: Wahrscheinlich (61-90% Chance), 3: Möglich (41-60% Chance), 2: Unwahrscheinlich (11-40% Chance), 1: Sehr unwahrscheinlich (weniger als 10% Chance).

5 Die Auswirkungen wurden anhand der folgenden Skala bewertet: 5: Katastrophal (irreversible Konsequenzen), 4: Kritisch (langfristige Konsequenzen), 3: Erheblich (signifikanter Schaden), 2: Geringfügig (lokalisierter oder minimaler Schaden), 1: Unbedeutend (keine bis kaum Auswirkungen).

Elektrizitätsgemeinschaften (LEG) sind räumliche Nähe und der Anschluss an die gleiche Netzebene eines Verteilnetzbetreibenden (ibid.).

Dies veranschaulicht das Community-Based Microgrid des NTC mit verteilten Energiequellen, welches auf den Arbeiten von Papaemmanouil et al. (2022, 2023) der Hochschule Luzern sowie dem NIST Framework for Smart Grid Interoperability Standards (2021) basiert.

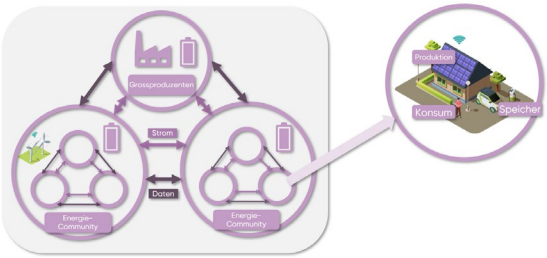


Abbildung 1 Energie-Community NTC

Rolle von Künstlicher Intelligenz (KI)

Künstliche Intelligenz wird in der Literaturrecherche als essenzielles Werkzeug für die Entwicklung benötigter digitaler Dienstleistungen im Stromökosystem (Barahona G. B., Bowler, Gounden, & Papaemmanouil, 2023, S. 75-82) und für die ökologisch und ökonomisch relevanten präzisen Vorhersagen von Stromproduktion und -verbrauch angesehen (Kern, et al., 2022).

Es ist davon auszugehen, dass die Verwaltung aller Teilnehmenden in einem Smart-Grid zu einem stark erhöhten Kommunikationsbedarf innerhalb der Strominfrastruktur führt. Die Bewältigung des zunehmenden Kommunikationsbedarfs wird für Menschen ohne den Einsatz von KI nicht mehr zuverlässig in Echtzeit bewältigt werden können. Dies gilt insbesondere bei der datengetriebenen Optimierung, beispielsweise des Verbrauchs und der Ladezyklen von Batterien. Aufgrund der Komplexität von LEG und der Notwendigkeit zur Echtzeit-Optimierung, dürfte der Bedarf von KI in LEG besonders hoch sein.

Das Szenario prognostiziert folglich, dass der Einsatz und die Abhängigkeit von KI-Systemen in intelligenten Stromnetzen zunehmen wird.

5 Hauptrisiken

Im Rahmen der Analyse wurden die folgenden acht Risiken identifiziert und als Grundlage für die Verifizierung durch Fachleute verwendet. Insgesamt wurden zwölf Interviews mit Expertinnen und Experten aus Zivilgesellschaft, Industrie und öffentlicher Verwaltung geführt. Sie wurden zu den identifizierten Risiken befragt und bewerteten jeweils die Eintrittswahrscheinlichkeit und die Schwere der potenziellen Auswirkungen auf das Schweizer Stromökosystem. Dabei wurde das Residualrisiko unter der Annahme beurteilt, dass das Stromsystem fachmännisch aufgebaut wird.

Nachfolgend wird auf die identifizierten Risiken sowie die Einschätzungen der Expertinnen und Experten eingegangen.

Es sei dabei angemerkt, dass diese Risiken nicht unbedingt neu sind. Neu ist jedoch der Einfluss der KI, von dem erwartet wird, dass er die Risiken verstärkt oder abschwächt. Sie müssen daher neu bewertet werden. Die Risiken wurden in der Reihenfolge ihrer Bewertung geordnet. Das am höchsten bewertete Risiko wird zuerst behandelt.

1. **Erhöhte Angriffsfläche kommerzieller Endkundenprodukte:** Kommerzielle Endkundenprodukte sind Produkte, die für den direkten Verkauf an Verbrauchende (Endkundinnen und -kunden) bestimmt sind. Diese Produkte werden für den persönlichen oder häuslichen Gebrauch entwickelt und häufig in grossen Stückzahlen vertrieben (im Gegensatz zu Produkten, die für Unternehmen oder den industriellen Gebrauch bestimmt sind). Komplexe und autonome Systeme, deren Funktionsweise durch die Inbetriebnehmenden und Anwendenden oft nur teilweise verstanden wird, finden sowohl in der Strominfrastruktur als auch in Endverbraucherprodukten für Unternehmen und Haushalte Anwendung. Insbesondere im Endkundenbereich bergen solche Systeme potenzielle Risiken, da ihr Markteintritt häufig durch kommerzielle Interessen beschleunigt wird, ohne dass die Systeme umfassend getestet oder vollständig verstanden sind. Durch unzureichende Cybersicherheit, falsche Annahmen über die Funktionsweise und fehlende Sicherheitsupdates bei kommerziellen, KI-basierten Produkten könnten Angriffe das Stromnetz beschädigen.

   Beispielsweise könnte eine gleichzeitige Aktivierung aller Wasserkocher, Backöfen und Kochfelder in einem Quartier (LEG) zu einem lokalen Stromausfall (Blackout) führen.

Einschätzung der Expertinnen und Experten:

Am höchsten bewertet haben die Expertinnen und Experten das Risiko von kommerziellen Endprodukten, die direkt oder indirekt in das Kommunikationsnetz des Stromsystems eingebunden sind. 11 von 12 der Befragten haben sich für eine der beiden höchsten Wahrscheinlichkeiten («wahrscheinlich» bis «sehr wahrscheinlich») ausgesprochen und die Auswirkungen als «erheblich» bis «kritisch» eingestuft. Die Expertinnen und Experten stimmten zu, dass Geräte, wie z.B. intelligente Haushaltsgeräte, aufgrund des Zeitdrucks für eine schnelle Markteinführung oft nicht ausreichend auf ihre Cybersicherheit getestet werden. Schwachstellen könnten ausgenutzt werden, um das Stromnetz zu destabilisieren, falsche Informationen einzuspeisen oder destruktive Verbrauchs- und Produktionsmuster zu erzeugen.

2. **Kaskadeneffekte aufgrund hoher Komplexität:** Obwohl Systeme auf der Infrastrukturseite in der Regel besser kontrolliert und analysiert werden können als bei Endverbraucherprodukten, besteht das Risiko, dass das Zusammenspiel verschiedener, miteinander vernetzter Komponenten nicht vollständig nachvollzogen wird. Dies kann insbesondere in Szenarien mit dezentralisierten Energiequellen oder dynamischen Netzanforderungen zu unvorhergesehenen Problemen führen. Auch könnten Falschinformationen durch manipulierte Smart Meter zu fehlerhaften Regeleingriffen führen.

   Beispiele dafür sind Kaskadeneffekte durch KI-Entscheidungen. Sie können zu Systemausfällen führen, wenn KI-Systeme verschiedener Energie-Communities nicht synchron arbeiten und selbstverstärkende Kaskaden auslösen («Hochschaukeln»). Diese Effekte entstehen, weil die einzelnen Systeme versuchen, sich basierend auf Bedarf und Preissignalen eigenständig zu optimieren. Wenn viele dieser Systeme gleichzeitig agieren, kann das Zusammenspiel sehr komplex werden. Dadurch entstehen unerwartete Wechselwirkungen, die das Stromnetz belasten oder überlasten können.

   Beispielsweise ist denkbar, dass ein aufwiegelnder Fall das Netz überlastet, weil sich jede Energie-Community basierend auf dem eigenen Bedarf und dem Preissignal selbst optimieren will und dies mit entsprechender Geschwindigkeit tut. Gleichzeitig getroffene, individuell rationale Zuschaltentscheidungen von Produzenten können das Netz überfordern und destabilisieren.

   Einschätzung der Expertinnen und Experten:

   Am zweithöchsten wurde das Risiko von Kaskadeneffekten ausgehend von der Komplexität des Zusammenspiels vieler KI-Systeme bewertet. Ein Eintreten sei «möglich» bis «wahrscheinlich», die Auswirkungen «kritisch». Die Expertinnen und Experten sind sich einig, dass eine isolierte Betrachtung einzelner Systeme zu kurz greift, da einige der risikobehafteten Effekte erst durch Kaskaden vieler unter Umständen richtiger Einzelentscheidungen entstehen. Es ist zu erwarten, dass die Komplexität aufgrund vom KI-Einsatz zunehmen wird.

3. **Unabsichtliche KI-Fehler:** Grobe, unabsichtliche Fehler individueller, aber wichtiger KI-Systeme könnten zu Systemausfällen führen.

   Beispielsweise könnten unzureichend getestete KI-Updates von Herstellern zu breiten Ausfällen führen.

   Einschätzung der Expertinnen und Experten:

   Grobe Fehler individueller, aber wichtiger KI-Systeme, welche zum Totalausfall führen könnten, wurden mit der Eintrittswahrscheinlichkeit «möglich» mit «kritischen» Folgen eingestuft.

   «Menschliches Versagen oder fehlendes Wissen der Nutzenden führen oft zu Systemfehlern.»
   Martin Steiger, Rechtsanwalt für Recht im digitalen Raum, Steiger Legal AG

   Betont wurde auch die Rolle des Menschen beim Betrieb von Stromsystemen. Mehrere Expertinnen und Experten wiesen auf die Risiken hin, die durch menschliches Versagen oder mangelnde Übersicht in kritischen Situationen entstehen können. Insbesondere in Krisensituationen könnten Betreiber aufgrund der Komplexität und der Fülle an Informationen falsche Entscheidungen treffen.

4. **Abhängigkeiten zu einzelnen Herstellern**: Während bei den Herstellern von Endgeräten und folglich bei den Konsumierenden, Produzierenden und den Prosumern eine grosse Inhomogenität herrscht, ist die Anzahl der Anbieter und Hersteller von KI-Modellen vergleichsweise klein. Die Konzentration auf wenige Anbieter kann das Stromökosystem fehleranfälliger machen und birgt das Risiko bedrohter Verfügbarkeit und Integrität der KI-Modelle.

   Dezentrale Energieressourcen (DER) bringen insofern neue Absicherungen, dass ein einzelner Ausfall eines Produzenten keine grösseren Reaktionen benötigt. Trotzdem birgt die Inhomogenität der Energie-Communities auch neue Gefahren. Beispielsweise könnten synchron getaktete Angriffe auf die KI von Prosumern (z.B. durch gleichzeitiges Konsumieren oder Produzieren) diese schädigen und zum Abwurf einer oder sämtlicher Energie Communities führen. Dies könnte für die Konsumierenden innerhalb dieser Communities einem Totalausfall gleichkommen.

   Beispielsweise können Hersteller die Abhängigkeiten gezielt ausnutzen. Aufgrund geopolitischer Differenzen könnten sie auf Updates verzichten oder einzelne Modelle gezielt unbrauchbar machen (beispielsweise sämtliche Wechselrichter eines Herstellers von Photovoltaikanlagen).

Einschätzung der Expertinnen und Experten:

Die Einschränkung der Verfügbarkeit und Integrität von KI-Modellen aufgrund absichtlicher Ausnutzung von Abhängigkeiten sei «möglich», mit potenziell «kritischen» Konsequenzen.

Besonders kritisch wurde das Klumpenrisiko betrachtet: Viele Expertinnen und Experten äusserten die Sorge, dass sich die Lieferketten für KI-Modelle (insbesondere für Large Language Models, kurz LLM) auf einige wenige Anbieter konzentrieren. Damit steigt das Risiko, dass bei Ausfällen oder Angriffen auf die Lieferkette das gesamte System in Mitleidenschaft gezogen wird.

«Der Ausfall eines Herstellers wäre ein ernstzunehmendes wirtschaftliches Risiko.»
Lukas Mäder, Redaktor Technologie, NZZ

Es wurde auch auf das Risiko hingewiesen, dass wichtige Lieferanten von Infrastrukturkomponenten und Endprodukten aus wirtschaftlichen Gründen Produktreihen einstellen könnten, die dann keine Sicherheitsupdates mehr erhalten, aber potenziell Jahrzehnte weiter betrieben werden (müssten). Viele Expertinnen und Experten bestätigen daher, dass man bei der Auswahl der jeweiligen Hersteller vorsichtiger sein und auf verdächtige oder seltsam anmutende Systeme eher verzichten werde. Open Source Software stellt nach Ansicht einiger Expertinnen und Experten eine mögliche Alternative dar, um Abhängigkeitsrisiken zu reduzieren.

Es wurde auch angemerkt, dass Abhängigkeitsrisiken im Zuge der aktuellen geopolitischen Veränderungen zunehmen und an Relevanz gewinnen. Das Risiko müsse vor dem Hintergrund der weltpolitischen Entwicklungen bewertet werden. Die Systemabhängigkeiten in Europa wurden kürzlich durch das NTC analysiert[6].

5. **Abfluss und Missbrauch sensitiver Daten:** Daten, die KI-Systemen zur Verfügung gestellt werden, könnten entweder vom Hersteller gesammelt oder durch Hacking abgegriffen werden, was zu ernsthaften Sicherheitsbedrohungen führen kann.

   Der wesentliche Unterschied zu Daten in traditionellen Systemen besteht darin, dass Nutzende und Betreibende sensible Daten eher einem KI-System anvertrauen, weil sie sich davon einen höheren Nutzen versprechen – getreu dem Motto, dass KI dann besonders gut funktioniert, wenn sie mit möglichst vielen Datenpunkten gefüttert wird.

   Beispielsweise könnten Echtzeit-Stromverbrauchsdaten genutzt werden, um leerstehende Gebäude zu identifizieren – beispielsweise zur Planung von Einbrüchen oder Diebstählen.

Einschätzung der Expertinnen und Experten:

Interessant ist, dass die Expertinnen und Experten dem Abfluss sensibler Daten zwar die zweithöchste Eintrittswahrscheinlichkeit zuschreiben («wahrscheinlich»), die systemischen Auswirkungen jedoch im Vergleich zu den anderen Risiken am tiefsten einschätzen («geringfügig» bis «erheblich»). Dennoch betonten einige Expertinnen und Experten die Bedeutung des Schutzes von Nutzerdaten als elementar.

6. **Störung bei der Datenübertragung**: Eine Überlastung oder Störung der Kommunikationssysteme könnte dazu führen, dass KI-Modelle keine oder nicht repräsentative Daten erhalten, was in automatisierten Fehlentscheidungen resultieren kann.

   Beispielsweise könnten klassische Denial-of-Service (DoS) Angriffe auf das Kommunikationsnetz oder das Durchtrennen von (mehreren) Kommunikationsleitungen zu Störungen oder Überlastungen der Kommunikationssysteme führen.

Einschätzung der Expertinnen und Experten:

Der Störung oder Überlastung der Datenkommunikationskanäle in einem Ausmass, dass KI-Modelle nicht-repräsentative oder keine Daten erhalten, wird die Eintrittswahrscheinlichkeit («möglich» bis «wahrscheinlich») bei «erheblichen» potenziellen Auswirkungen zugeschrieben.

7. **Backdoors und gezielte Manipulation:** In den KI-Modellen eingebaute Hintertüren («backdoors») erlauben gezielte Manipulationen der Energie-Infrastruktur bis hin zum totalen Ausfall («kill switch»). Diese Backdoors könnten nicht nur vom Hersteller stammen, sondern auch von den Datenlieferanten, die für das Training der Modelle verantwortlich sind. Das Risiko ist stark abhängig von geopolitischen Einflüssen.

   Backdoors könnten ausgenutzt werden, um den lokalen Stromhandel zu manipulieren. So könnte beispielsweise eine gezielte destruktive Überlastung von Stromleitungen als Teil einer geostrategischen Cyberattacke genutzt werden, um Stromleitungen oder elektrische Komponenten aufgrund von über-

[6] Kurzbericht Systemabhängigkeiten Europas (NTC, 2024): https://www.ntc.swiss/news/2024-kurzbericht-systemabhaengigkeiten

mässiger Strombelastung zu beschädigen oder zu zerstören.

Einschätzung der Expertinnen und Experten:

Die Expertinnen und Experten bestätigen die Tragweite des Risikos von Hintertüren.

Obschon sie Backdoors «nur» als «möglich» einstufen, schätzen sie potenzielle Auswirkungen als «kritisch» ein.

Besonders brisant wird dieses Risiko durch geopolitische Einflüsse: Die Systeme werden immer komplexer, und wo es keinen freien Markt mit Hunderten von Anbietern gibt, lässt sich der Einbau von Hintertüren aus geopolitischen Gründen vermutlich nicht verhindern. Mehrere Expertinnen und Experten wiesen darauf hin, dass die geopolitische Lage einen erheblichen Einfluss auf die Risikoeinschätzung hat. Während diese Analyse in einer unangespannten Lage der Schweiz durchgeführt wurde, müsste beispielsweise das Sabotagerisiko in Ausnahmesituationen wie geopolitischen Spannungen oder Konflikten deutlich höher eingeschätzt werden.

«Die Risikoeinschätzung ist stark abhängig von der geopolitischen Lage. Das Stromökosystem ist in Ausnahmesituationen sehr anfällig für Sabotageakte.»
Markus Rössli, Direktor PTI Schweiz

KI wird zunehmend Hardware nutzen, welche dieselben Black-Box Risiken beinhalten wie beispielsweise Central Processing Units (CPU). Während softwarebasierte Hintertüren in der Regel schwer zu verbergen sind, betonten die Expertinnen und Experten, dass hardwarebasierte Hintertüren eine weitaus realistischere und zugleich gefährlichere Bedrohung darstellen. Diese Schwachstellen sind oft schwerer zu entdecken und können für gezielte Manipulationen ausgenutzt werden.

«Hardwarebasierte Schwachstellen sind viel schwieriger zu entdecken und sehr realistisch, sie werden bereits heute für gezielte Manipulationen ausgenutzt.»
Florian Tramèr, Assistant Professor of Computer Science, ETH Zürich

8. **Absichtliche Manipulation der Datengrundlage:** KI-Systeme sind von ausreichender Datenqualität abhängig, da dezentrale KI-Modelle auf deren Basis Prognosen generieren und Entscheidungen treffen. Während diese Entwicklung zu mehr Effizienz und Optimierung führt, könnte ein Ausfall oder eine Störung des Datenaustauschs oder der Entscheidungssysteme weitreichende Folgen haben. Eine gezielte Manipulation von Daten, könnte falsche Entscheidungen auslösen und zu Ausfällen führen.

   Beispielsweise könnten Manipulationen von Wettervorhersagen dazu führen, dass fälschlicherweise eine besonders hohe Sonneneinstrahlung erwartet wird, worauf die Betreibenden ihre Batterien vorzeitig entladen, um später günstigen Solarstrom aus der erwarteten Überproduktion der Photovoltaikanlagen (PV) wieder aufzuladen. Die daraus resultierende Fehlplanung könnte zu einem Mangel an Regelstrom und in der Folge zu Ausfällen führen. Mit der Einspeisung falscher Verbrauchsinformationen ins Kommunikationsnetzwerk, könnten ausserdem fehlgeleitete Energieplanungen, ungenaue Lastverteilungen, fehlerhafte Preisgestaltungen und Beeinträchtigungen der Netzstabilität herbeigeführt werden.

Einschätzung der Expertinnen und Experten:

Obschon dieses Risiko im Vergleich mit anderen Risiken am tiefsten eingestuft wurde, wird die Eintrittswahrscheinlichkeit dennoch als «möglich», die Auswirkungen als «erheblich bis kritisch» eingestuft.

6 Massnahmen und Empfehlungen

Die Ergebnisse unterstreichen die Notwendigkeit einer proaktiven und intensiven Auseinandersetzung mit den durch KI herbeigeführten Herausforderungen durch eine neutrale Organisation, beispielsweise durch das NTC. Insbesondere im Hinblick auf die Entwicklung des Schweizer Stromökosystems hin zu intelligenten Netzen (Smart Grids) und die wachsende Bedeutung von KI im Stromökosystem ist eine stärkere Sensibilisierung und laufende Prüfung der Regulierungen notwendig, um zukünftigen Bedrohungen angemessen begegnen zu können.

Die Auswertung verdeutlicht insbesondere, dass ein grosses Risiko von kommerziellen Endprodukten ausgeht, die von Betreibern und Haushalten verwendet werden und an der Kommunikation im Stromnetz beteiligt sind. Eine koordinierte und wiederholte proaktive Identifizierung von Schwachstellen durch Cybersicherheitsprüfungen wird dringend empfohlen.

Ein weiteres wesentliches Risiko liegt in der Komplexität des Zusammenspiels vieler KI-Systeme. Betreibende und Nutzende von KI-Systemen müssen hinsichtlich einer holistischen Betrachtungsweise sensibilisiert werden. Eine isolierte Betrachtung einzelner Systeme greift zu kurz, da einige der risikobehafteten Effekte erst durch Kaskaden vieler unter Umständen richtiger Einzelentscheidungen entstehen. Regelmässige Schulungen für Anwendende und Betreibende sind unerlässlich.

Durch eine diversifizierte Beschaffungsstrategie sollen zudem Abhängigkeiten der wenigen KI-Modell-Anbieter und geopolitische Risiken reduziert werden.

Die Risiken werden mit der Komplexität der Automatisierung weiter steigen. Es empfiehlt sich daher, die Entwicklungen genau zu beobachten und entsprechende Infrastrukturen mit den Mitteln auszustatten, um auf diese Risiken zu reagieren.

Laut einiger Expertinnen und Experten sei die Bereitschaft der Energiebranche, sich intensiv mit technischen und regulatorischen Vorgaben auseinanderzusetzen, eher gering. Dies könnte die Umsetzung von Sicherheitsmassnahmen und die Anpassung an neue Risiken verlangsamen. Dies komplett dem freien Markt zu überlassen, könnte zu unbefriedigenden Cybersicherheitskonstellationen und Abhängigkeiten führen.

Positiv ist, dass das Schweizer Verteilnetz aktuell und auch in Zukunft von Fachspezialistinnen und -spezialisten entwickelt und betrieben wird. Es sind logische, physische Sicherungen und Regelsysteme im Einsatz, die das Stromnetz im Normalfall absichern. Die Expertinnen und Experten hoben diesbezüglich hervor, dass die absehbare Abhängigkeit von automatisierten Informations- und Entscheidungsprozessen die Bedeutung konventioneller Sicherungssysteme unterstreiche.

Konventionelle Schutzmassnahmen funktionieren unabhängig von digitalen oder automatisierten Systemen. Bei einem Ausfall oder Cyberangriff können solche physischen Sicherungssysteme weiterhin Schutz bieten und verhindern, dass der Ausfall eskaliert. So betonten einige Expertinnen und Experten die Relevanz physischer Schutzmassnahmen, wie beispielsweise thermische Sicherungen, um das Netz bei Ausfällen oder Cyber-Angriffen zu schützen. Dies zeigte sich auch in der Diskussion über frühere Netzstörungen. Einige Expertinnen und Experten verwiesen auf den Vorfall vom 23. September 2003, als in Italien ein Baum auf eine 380-kV-Leitung fiel und einen grossflächigen Stromausfall verursachte (UCTE, 2003). Konventionelle Sicherungen trugen dazu bei, die Auswirkungen zu kontrollieren und weitere Schäden zu verhindern.

7 Zusammenfassung

Es besteht Konsens über die Relevanz und Dringlichkeit der identifizierten Top-Risiken. Die Expertinnen und Experten haben in den Gesprächen deutlich die Wichtigkeit dieses Themas gelobt und bekräftigt, dass es sich um ernstzunehmende Risiken handelt, die berücksichtigt werden müssen.

Auffallend ist vor allem die eher hohe Risikoeinschätzung der Expertinnen und Experten. Viele Risiken wurden als „hoch" oder „sehr hoch" eingestuft. Die Kategorie „extrem" wurde nur selten vergeben. Trotz des hohen Abstraktionsgrades des Zukunftsszenarios waren die Antworten von den bisherigen Erfahrungen mit Cybervorfällen geprägt, die in vielen Fällen aufgrund mangelnder Ressourcen auf unzureichend gesicherte Infrastrukturen zurückzuführen waren.

Die Untersuchung hat gezeigt, dass der Einsatz von KI im Schweizer Stromökosystem neben zahlreichen Vorteilen (z.B. Effizienzsteigerung) auch erhebliche Risiken mit sich bringt.

Den Risiken muss mit der notwendigen Sensibilisierung und Regulierung, den notwendigen Mitteln, sowie regelmässigen Cybersicherheitsprüfungen entgegnet werden.

Viele der identifizierten Risiken und Argumente sind nicht spezifisch für die Strominfrastruktur und lassen sich auch auf andere Anwendungsbereiche von KI übertragen. Damit kann diese Arbeit des NTC generisch verstanden werden.

8 Limitationen

Der vorliegende Bericht zeigt auf, welche Risiken im Zusammenhang mit KI bestehen und wie diese zu gewichten sind. Allerdings gibt der Bericht nicht Aufschluss darüber, wie stark sich die Risiken durch KI im Vergleich zu klassischen IT-Systemen verändern. Die Expertinnen und Experten haben berechtigterweise angemerkt, dass die Abgrenzung von Künstlicher Intelligenz (KI) zu klassischen IT-Systemen bei der Definition der Risikokategorien sowie deren spezifischer Anwendungen in der Strominfrastruktur optimiert werden könnte. Dies ist in vorliegendem Bericht zum Teil auf den bewusst hohen Abstraktionsgrad des Szenarios zurückzuführen. Künftige Recherchen sollten eine differenzierte Abgrenzung berücksichtigen.

Der Bericht basiert auf einem möglichen Zukunftsszenario mit hohem Abstraktionsgrad. Expertinnen und Experten aus den Bereichen Cybersicherheit, Digitalisierung und Künstliche Intelligenz KI haben ihre Einschätzungen zu möglichen KI-Risiken abgegeben. Daher handelt es sich bei dem Bericht nicht um eine wissenschaftliche Studie, die mit Expertinnen und Experten aus der Stromwirtschaft durchgeführt wurde und auf einem realen KI-basierten Stromökosystem basiert.

Die vorliegende Risikoeinschätzung ist eine Momentaufnahme und muss laufend reevaluiert werden (insbesondere bei sich verändernder geopolitischer Lage).

9 Anhang

Abkürzungsverzeichnis

NTC
Nationales Testinstitut für Cybersicherheit NTC

KI
Künstliche Intelligenz

LEG
Lokale Elektrizitätsgemeinschaften

DER
Distributed Energy Resources

DoS
Denial of Service

Literaturverzeichnis

Barahona, G. B., Bowler, B., Gounden, C., & Papaemmanouil, A. (2023). *Business Models of an AI Marketplace for Energy Systems with Focus on Demand Response.* (Bde. Smart Services Summit. Smart Services Creating Sustainability, 2023,). doi:https://doi.org/10.5281/zenodo.11126792

BFE. (2021). *Intelligente Netze (Smart Grids) und intelligente Messsysteme (Smart Metering).* Von https://www.bfe.admin.ch/bfe/de/home/versorgung/stromversorgung/stromnetze/smart-grids.html abgerufen

BFE. (2024). *Vorlage für eine sichere Stromversorgung.* Von https://www.bfe.admin.ch/bfe/de/home/versorgung/stromversorgung/bundesgesetz-erneuerbare-stromversorgung.html abgerufen

Blick. (2024). Aprilwetter sorgte beinahe für Blackout. *Blick*. Von https://www.blick.ch/wirtschaft/stromproduzenten-verkalkulieren-sich-komplett-wintereinbruch-kostet-konsumenten-30-millionen-franken-an-strom-id19692455.html abgerufen

International Energy Agency. (n.d.). *Smart Grids.* Von https://www.iea.org/energy-system/electricity/smart-grids abgerufen

Kern, D., Ensinger, A., Hammer, C., Neufeld, C., Lecon, C., Nagl, A., . . . Wood, B. M. (2022). *Application Possibilities of Artificial Intelligence in a Renewable Energy Platform.* (Bd. Smart Services Summit. Progress in IS.). Springer, Cham. doi:https://doi.org/10.1007/978-3-030-97042-0_

NIST. (2021). *NIST Framework and Roadmap for Smart Grid Interoperability Standardss, Release 4.0.* doi:https://doi.org/10.6028/NIST.SP.1108r4

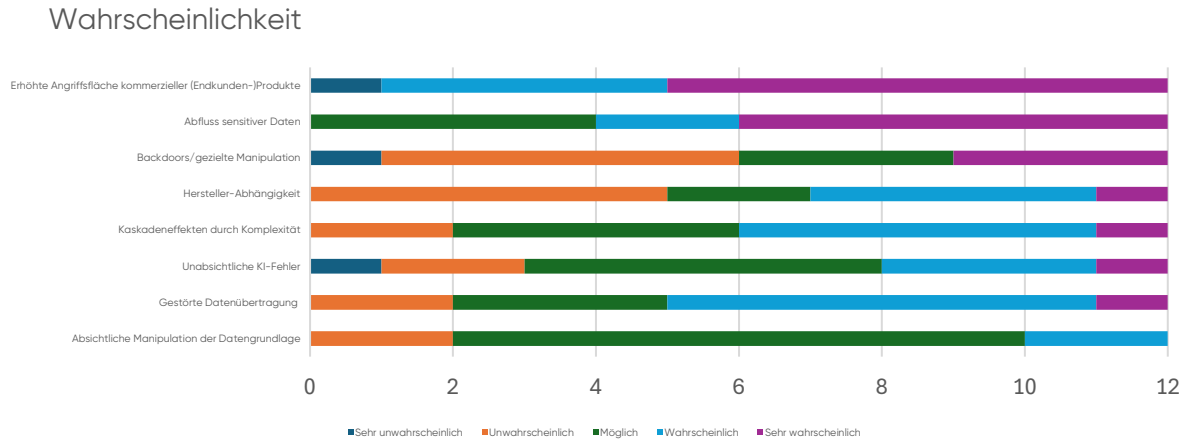
Trivedi, R., Patra, S., Sidqi, Y., Bowler, B., Zimmermann, F., Deconinck, G., . . . Khadem, S. (2022). *Community-Based Microgrids: Literature Review and Pathways to Decarbonise the Local Electricity Network.* doi:https://doi.org/10.3390/en15030918

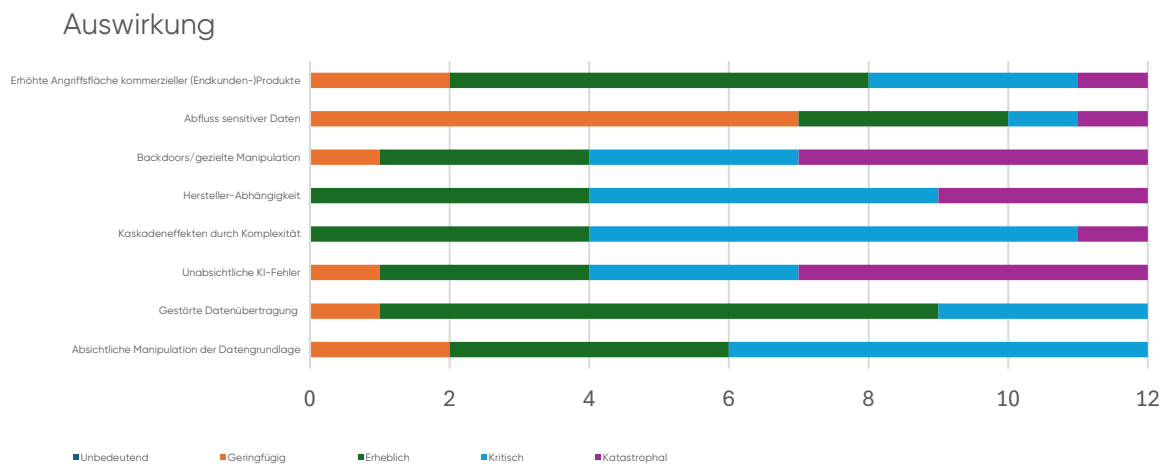
UCTE. (2003). *FINAL REPORT of the INvestigation Committee on the 28. September 2003 Blackout in Italy.* Von https://eepublicdownloads.entsoe.eu/clean-documents/pre2015/publications/ce/otherreports/20040427_UCTE_IC_Final_report.pdf abgerufen

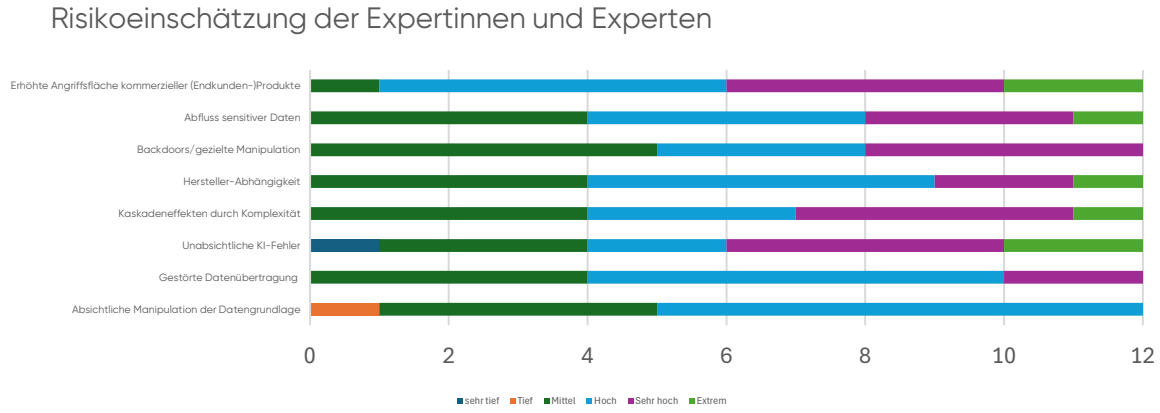
UVEK. (2024). *Bundesgesetz über eine sichere Stromversorgung mit erneuerbaren Energien: Änderung der Stromversorgungsverordnung | Erläuternder Bericht zur Vernehmlassungsvorlage.* Von https://pubdb.bfe.admin.ch/de/publication/download/11641 abgerufen

VSE. (2022). *Die Energieversorgung der Schweiz im Jahr 2050.* Von https://www.strom.ch/de/energiezukunft-2050/resultate abgerufen

Grafiken


Wahrscheinlichkeit
Erhöhte Angriffsfläche kommerzieller (Endkunden-)Produkte
Abfluss sensitiver Daten
Backdoors/gezielte Manipulation
Hersteller-Abhängigkeit
Kaskadeneffekten durch Komplexität
Unabsichtliche KI-Fehler
Gestörte Datenübertragung
Absichtliche Manipulation der Datengrundlage
0
2
4
6
8
10
12
Sehr unwahrscheinlich
Unwahrscheinlich
Möglich
Wahrscheinlich
Sehr wahrscheinlich


Auswirkung
Erhöhte Angriffsfläche kommerzieller (Endkunden-)Produkte
Abfluss sensitiver Daten
Backdoors/gezielte Manipulation
Hersteller-Abhängigkeit
Kaskadeneffekten durch Komplexität
Unabsichtliche KI-Fehler
Gestörte Datenübertragung
Absichtliche Manipulation der Datengrundlage
0
2
4
6
8
10
12
Unbedeutend
Geringfügig
Erheblich
Kritisch
Katastrophal


Risikoeinschätzung der Expertinnen und Experten
Erhöhte Angriffsfläche kommerzieller (Endkunden-)Produkte
Abfluss sensitiver Daten
Backdoors/gezielte Manipulation
Hersteller-Abhängigkeit
Kaskadeneffekten durch Komplexität
Unabsichtliche KI-Fehler
Gestörte Datenübertragung
Absichtliche Manipulation der Datengrundlage
0
2
4
6
8
10
12
sehr tief
Tief
Mittel
Hoch
Sehr hoch
Extrem