

SSC Report

# Synergise. Strategise. Realise.

SSC recommendations for AI computing infrastructure in the ERI domain



Schweizerische Eidgenossenschaft  
Confédération suisse  
Confederazione Svizzera  
Confederaziun svizra

Schweizerischer Wissenschaftsrat SWR  
Conseil suisse de la science CSS  
Consiglio svizzero della scienza CSS  
Swiss Science Council SSC



## Der Schweizerische Wissenschaftsrat

Der Schweizerische Wissenschaftsrat SWR berät den Bund in allen Fragen der Wissenschafts-, Hochschul-, Forschungs- und Innovationspolitik. Ziel seiner Arbeit ist die kontinuierliche Optimierung der Rahmenbedingungen für die gedeihliche Entwicklung der Schweizer Bildungs-, Forschungs- und Innovationslandschaft. Als unabhängiges Beratungsorgan des Bundesrates nimmt der SWR eine Langzeitperspektive auf das gesamte BFI-System ein.

## Le Conseil suisse de la science

Le Conseil suisse de la science CSS est l'organe consultatif du Conseil fédéral pour les questions relevant de la politique de la science, des hautes écoles, de la recherche et de l'innovation. Le but de son travail est l'amélioration constante des conditions-cadre de l'espace suisse de la formation, de la recherche et de l'innovation en vue de son développement optimal. En tant qu'organe consultatif indépendant, le CSS prend position dans une perspective à long terme sur le système suisse de formation, de recherche et d'innovation.

## Il Consiglio svizzero della scienza

Il Consiglio svizzero della scienza CSS è l'organo consultivo del Consiglio federale per le questioni riguardanti la politica in materia di scienza, scuole universitarie, ricerca e innovazione. L'obiettivo del suo lavoro è migliorare le condizioni quadro per lo spazio svizzero della formazione, della ricerca e dell'innovazione affinché possa svilupparsi in modo armonioso. In qualità di organo consultivo indipendente del Consiglio federale, il CSS guarda al sistema svizzero della formazione, della ricerca e dell'innovazione in una prospettiva globale e a lungo termine.

## The Swiss Science Council

The Swiss Science Council SSC is the advisory body to the Federal Council for issues related to science, higher education, research and innovation policy. The goal of the SSC, in conformity with its role as an independent consultative body, is to promote the framework for the successful development of the Swiss higher education, research and innovation system. As an independent advisory body to the Federal Council, the SSC pursues the Swiss higher education, research and innovation landscape from a long-term perspective.

[www.wissenschaftsrat.ch](http://www.wissenschaftsrat.ch)

# Content

<b>Management Summary</b>	<b>6</b>
Recommendations	7
<b>Zusammenfassung</b>	<b>11</b>
Empfehlungen	12
<b>Résumé</b>	<b>18</b>
Recommandations	20
<b>Riassunto</b>	<b>24</b>
Raccomandazioni	25
<b>1 Framing and context</b>	<b>30</b>
1.1 Why this topic – AI and computing infrastructure?	30
1.2 What currently exists in terms of AI and computing infrastructure	31
1.3 Approach to address AI and computing infrastructure	36
<b>2 Reasoning behind the recommendations</b>	<b>37</b>
2.1 National strategy – guiding principles	37
2.2 Tiered computing infrastructure system	39
2.3 Agile computing infrastructure strategy board	45
2.4 Financial implications	47
2.5 What are the benefits of implementing these recommendations?	47
<b>3 What other foundation stones are needed?</b>	<b>48</b>

<b>4</b>	<b>Annex</b>	<b>49</b>
4.1	National and international interview guides	49
4.2	Working process	55
4.3	Definitions	59
4.4	Abbreviations	60

## Tables of figures

Figure 1:		
Illustration of the tiered computing infrastructure system		9
Abbildung 2:		
Illustration des Systems der mehrstufigen Recheninfrastruktur		15
Figure 3 :		
Représentation schématique du système d'infrastructure de calcul à plusieurs niveaux.		22
Figura 4:		
Illustrazione di un sistema di infrastruttura di calcolo a più livelli.		28
Figure 5:		
Illustration of the tiered computing infrastructure system.		40

# Management Summary

Over the past decade, the importance of data-driven research has grown, supported by artificial intelligence (AI) methods such as machine learning. Adequate computing infrastructure is essential for data- and compute-intensive analyses. The Swiss Science Council (SSC) has examined whether Swiss academia has the necessary resources to meet its current and future needs. Through interviews with organisations in Switzerland and abroad, the SSC has gathered information on current computing infrastructures, cloud services and future plans and visions. Swiss higher education institutions (HEI) are actively considering how to adapt their systems to meet rising computational demands. This is considered strategically important for the future of the Swiss education, research and innovation (ERI) system, and several challenges have been identified.

Predicting resources is difficult because future computing needs and the effects of optimisation are difficult to forecast. HEI face challenges as their computing infrastructures are limited, become outdated quickly and are less competitive than those in industry. As industry increasingly leads AI research, academia risks falling behind and must find ways to remain competitive. Existing academic computing resources are often considered inadequate, prompting some experts to collaborate with industry to gain better access. The Swiss AI Initiative is perceived as benefiting the ETH domain more than it serves as a national initiative associated with the most advanced computing resources.

HEI emphasise the importance of planning data access, storage and archiving alongside computing infrastructure planning. Concerns about data privacy, protection, sovereignty and costs lead to a preference for in-house computing. Swiss natural resources are considered insufficient or too expensive for the kind of large-scale supercomputing seen internationally. Retaining

senior AI talent is challenging in a competitive job market, which has resulted in a brain drain from academia to industry.

Based on the results of the national and international interviews conducted, Switzerland's diverse academic landscape currently appears suboptimally prepared for future research requiring intensive computing resources, such as AI. However, the design of future-oriented, high-performance computing capacities is being addressed at political and strategic levels.<sup>1,2,3,4,5</sup> As this field is subject to rapid change and its importance for Switzerland cannot be doubted, the council recommends the development of a long-term AI infrastructure strategy and to create a tiered computing infrastructure system based on this national strategy. This requires establishing a national strategic board and securing adequate funding. For this to happen, the council recommends as **immediate action**: Declaring the development of a computing infrastructure for academic research a national task for which new mechanisms must be found, both in terms of design and financing, in order to master the challenges as identified. Make it accessible and equally usable for the diverse research, innovation and education landscape, in line with its needs and compatible with economic strategies, so that economic benefits would also be targeted, for example, through the integration and servicing of SME. To this end, the relevant stakeholders (e.g., higher education institutions, federation, cantons) must engage in dialogue in order to develop, coordinate, establish and finance the necessary mechanisms at national level with a long-term commitment.

- 1 Federal Chancellery FCh. Digital Switzerland Strategy 2025. January 10, 2025. BBl 2025 31. <https://www.fedlex.admin.ch/eli/fga/2025/31/de> and <https://digital.swiss/en/strategy/focus-topics/artificial-intelligence> (last checked on 26.01.26).
- 2 Federal Office of Communications OFCOM. Overview of the regulation of artificial intelligence. Report to the Federal Council. February 12, 2025. <https://www.bj.admin.ch/bj/de/home/staat/gesetzgebung/kuenstliche-intelligenz.html> (last checked on 26.01.26; report not available in English).
- 3 Federal Chancellery FCh. Strategy Use of AI systems in the Federal Administration. March 21, 2025. <https://www.bk.admin.ch/bk/en/home/digitale-transformation-ikt-lenkung/vorgaben/sb021-strategie-einsatz-von-ki-systemen-in-der-bundesverwaltung.html> (last checked on 26.01.26).
- 4 Swiss Federal Audit Office SFAO. International parallel audit on artificial intelligence. SFAO-25128. July 28, 2025. [https://www.efk.admin.ch/wp-content/uploads/publikationen/berichte/wirtschaft\\_und\\_verwaltung/informatikprojekte/25128/25128\\_endgueltige\\_fassung\\_v04.pdf](https://www.efk.admin.ch/wp-content/uploads/publikationen/berichte/wirtschaft_und_verwaltung/informatikprojekte/25128/25128_endgueltige_fassung_v04.pdf) (last checked on 26.01.26; report not available in English).
- 5 Federal Chancellery FCh. Further development of coordination for the use of artificial intelligence in the federal administration. Concept. December 12, 2025. <https://www.bk.admin.ch/dam/bk/en/dokumente/dti/themen/ki/konzept-weiterentwicklung-koordination-ki.pdf.download.pdf/konzept-weiterentwicklung-koordination-ki.pdf> (last checked on 26.01.26; report not available in English).

## Recommendations

### National strategy – guiding principles

Switzerland needs to develop a long-term AI infrastructure strategy that serves all public ERI needs. This national strategy has to include a modular computing infrastructure leveraging existing synergies, by forecasting and aligning the needs of the ERI domain's national large-scale computing resources and by enabling Switzerland to participate in the development of international computing infrastructures. It is to outline how a tiered computing infrastructure system meets user requirements for flexibility, scalability and efficiency, including cost-efficiency, while also being resilient and sustainable considering natural resources. It should value digital sovereignty and knowledge security, and foster partnerships with value-sharing stakeholders. It should include a data lifecycle management strategy for the tiered computing infrastructure system that builds on the FAIR<sup>6</sup> principles, and it also ought to set out how to promote and sustain world-class AI expertise in Switzerland that can compete internationally.

For the purpose of establishing such a national strategy, the SSC recommends that computing resources include all the necessary components, services and personnel required to make them usable for end users. Consequently, they require long-term commitment and investment. This is why the Council believes that enabling the use of AI for ERI stakeholders and other computationally intensive academic research requires appropriate computing infrastructure that leverages synergies and follows a national strategy. To achieve this, the Council believes that creating a computing infrastructure in the form of a tiered system is essential. Building such a system requires establishing a national strategic board and securing adequate funding.

<sup>6</sup> FAIR: Findable, accessible, interoperable, reusable. Wilkinson, M., Dumontier, M., Aalbersberg, I. et al. The FAIR Guiding Principles for scientific data management and stewardship. *Sci Data* 3, 160018 (2016). FAIR4RS. Barker, M., Chue Hong, N.P., Katz, D.S. et al. Introducing the FAIR Principles for research software. *Sci Data* 9, 622 (2022).

The Council recommends:

## 1. Tiered computing infrastructure system

- To build a tiered computing infrastructure system based on a national strategy and implemented by its stakeholders, guided by the principles outlined above.

For Tier 0 and 1<sup>7</sup>: International and national supercomputing<sup>8</sup>

- To further develop the existing tiers of national and international supercomputing by expanding the access to and participation in the highest-end international computing resources and integrating them with the national resources.
- To place the existing national supercomputing provider (CSCS), which also provides access to international supercomputing facilities<sup>9</sup>, under a governance that fulfils its national mandate, to ensure serving all ERI stakeholders throughout Switzerland.

For Tier 2: National high-performance computing

- To promote a national high-performance computing infrastructure which serves a wide variety of users who do not require supercomputing resources and which integrates or builds on existing computing resources.

For Tier 3: Regional compute centres<sup>10</sup>

- To set up local or regional AI expert groups at higher education institutions that support users locally and direct them to the appropriate computing resources within the tiered system, as required.
- The tiers should provide computing resources that include all the necessary components, services and resources required to make computing power usable by end users, thereby also functioning as pools of national experts for user and scientific support including AI developments.

This will provide Switzerland's diverse academic landscape with a comprehensive, appropriately scaled, sustainable and resilient computing infrastructure serving a wide range of needs. The SSC envisions a tiered computing infrastructure system that enables greater efficiency by building on interoperability and providing users with seamless transition between the tiers, e.g., for federated learning, and is managed through a single point of access (e.g., for tiers 0, 1 and 2), while taking into account the different funding sources and governance structures already in place.

<sup>7</sup> Example: CSCS: link to international supercomputing facility – LUMI (FIN); national supercomputing: Alps.

<sup>8</sup> In this context, 'supercomputing' refers to 'highest-end computing', the most advanced systems operating at the furthest limits of current technology.

<sup>9</sup> Access to LUMI: <https://www.lumi-supercomputer.eu/get-started-2021/users-in-switzerland/> and through the activities of HELvetic AI Resources, Technologies and Services (HEARTS), see <https://csc.fi/en/news/eurohpc-ju-se-lects-ai-factory-antennas-to-broaden-ai-factories-initiative/> and <https://goba.swiss/en/switzerland-joins-europe-an-ai-factory-network-with-new-hearts-antenna/> (last checked on 26.01.26).

<sup>10</sup> Example: HEI: central HPC (on-premise solutions) of cantonal universities or of UAS.

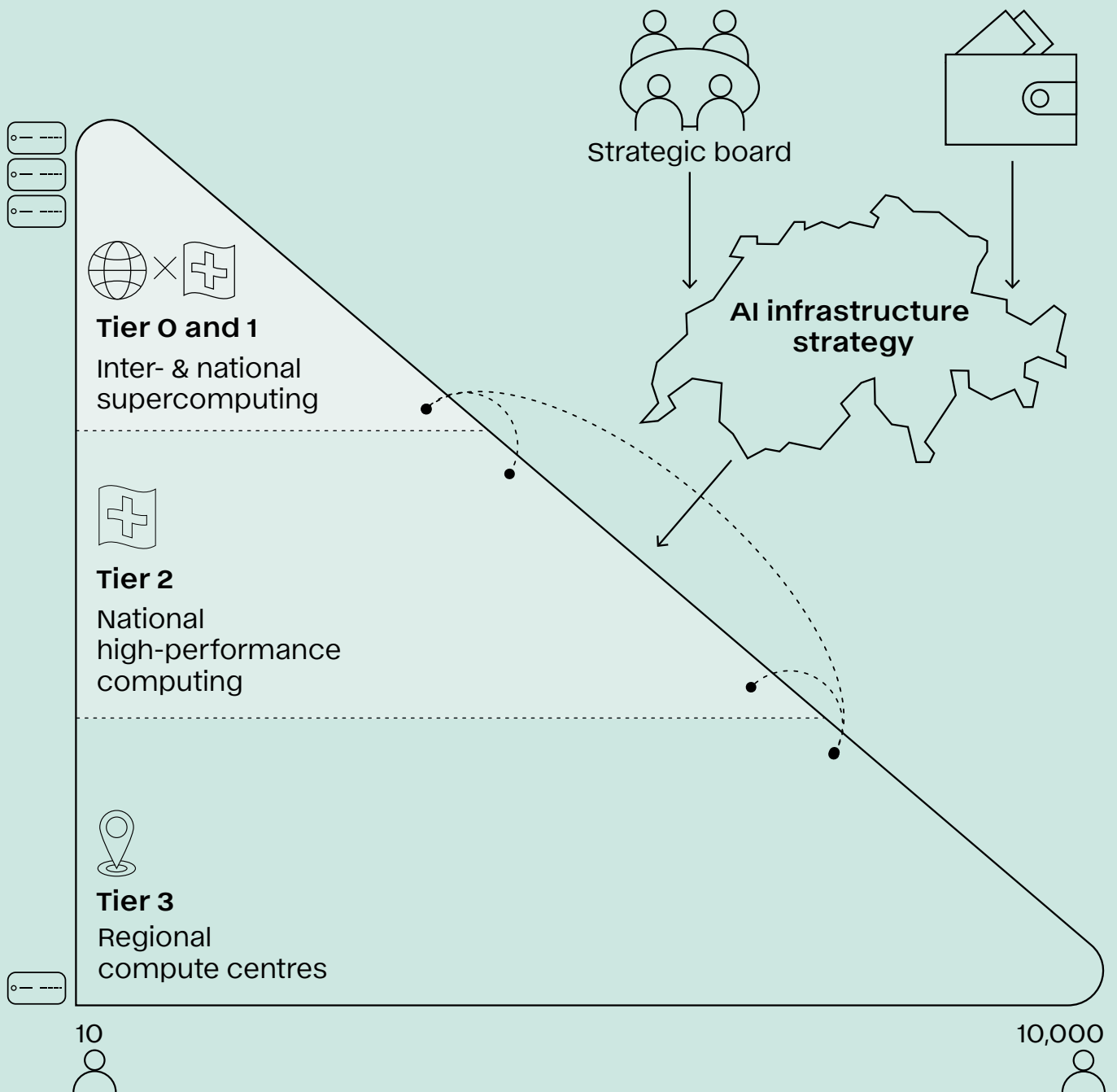


Figure 1: Illustration of the tiered computing infrastructure system.

The X-axis exemplifies the number of users of the system, while the Y-axis shows the level of computing capacity. The number of users shown on the X-axis does not represent actual user numbers. It is for illustrative purposes only. The dotted lines represent permeability between the individual tiers. The dotted semicircles represent interoperability. Interoperability here refers to the ability to exchange and use information across all levels, e.g., data (incl. software) and computing capacity.

## 2. Agile computing infrastructure strategy board

- To establish a national, independent and neutral strategic board that assesses and decides on the design of a future-oriented computing infrastructure in form of a tiered system. The strategic board is responsible for:
  - designing the tiered system strategically, based on its continuous assessment of national large-scale computing resource needs and international computing infrastructure developments, taking into account multilevel governance. The national strategy should be based on the guiding principles outlined above.
  - monitoring the implementation and performance of the national strategy by the single tiers of the tiered computing infrastructure system.

The Federal Council should appoint this strategic board as an independent commission of experts. It should be equipped with the necessary powers to fulfil its responsibilities, if necessary on a legal basis.

- The strategic board should comprise all relevant competencies required to undertake their accountabilities, including technical<sup>11</sup>, scientific<sup>12</sup>, strategic<sup>13</sup>, legal<sup>14</sup> and financial<sup>15</sup> expertise, simultaneously representing users and interest groups of national and regional computing resources of higher education institutions, with the size of the board allowing for efficient steering.

## 3. Financial implications

- To adapt funding in order to reimburse computing resources in their entirety, financing can only partially rely on grants, as these only provide for short-term costs. Computing resources, however, require long-term commitment and investment. Therefore, adequate long-term funding must be provided for all tiers. In general, computing resources should be provided in a transparent, cost-effective manner, meaning that costs must be covered but no profit should be made from higher education institutions.

11 Technical: advanced hardware architectures, accelerators and specialised chips, system software and programming, optimisation and performance tuning, network and data infrastructures, data management and emerging technologies (e.g., quantum computing).

12 Scientific: domain-specific research expertise: understanding the unique computational, data, and simulation needs of various areas of research to ensure that HPC system designs align with real-world scientific demands.

13 Strategic: interdisciplinary collaboration and strategic foresight: fostering collaboration among technical, scientific, policy and business experts, while maintaining a long-term, strategic vision to align the computing infrastructure with evolving research and market trends.

14 Legal: ecosystem and supply chain dynamics: understanding of supplier relationships, intellectual property considerations and compliance with regulatory frameworks to mitigate risks and ensure a legally sound computing infrastructure, including associated storage and archiving systems. Risk management and regulatory compliance: managing legal and cybersecurity risks, including adherence to export controls, data privacy regulations and other legal obligations, to ensure that strategic investments are both resilient and compliant.

15 Financial: strategic investment analysis and financial acumen: ability to conduct thorough cost-benefit analyses and manage long-term investment strategies, balancing performance, scalability and fiscal responsibility.

# Zusammenfassung

In den letzten zehn Jahren hat die datengetriebene Forschung kontinuierlich an Bedeutung gewonnen, nicht zuletzt durch den Einsatz von Methoden der künstlichen Intelligenz (KI) wie dem maschinellen Lernen. Für daten- und rechenintensive Analysen ist eine angemessene Recheninfrastruktur unerlässlich. Der Schweizerische Wissenschaftsrat (SWR) hat untersucht, ob die Schweizer Wissenschaft über die erforderlichen Ressourcen verfügt, um ihren aktuellen und zukünftigen Bedarf zu decken. Zu diesem Zweck interviewte er Organisationen in der Schweiz und im Ausland und analysierte Informationen zu bestehenden Recheninfrastrukturen, Cloud-Diensten sowie zukünftigen Plänen und Visionen. Die Schweizer Hochschulen befassen sich aktiv mit der Frage, wie sie ihre Systeme an die steigenden Anforderungen an Rechenkapazität anpassen können. Diese Frage gilt als strategisch bedeutsam für die Zukunft des schweizerischen Bildungs-, Forschungs- und Innovationssystems (BFI); entsprechend wurden mehrere Herausforderungen identifiziert.

Die Ressourcenplanung gestaltet sich schwierig, da sowohl der künftige Bedarf an Rechenkapazität als auch die Auswirkungen technischer Optimierungen schwer abschätzbar sind. Hochschulen stehen vor Herausforderungen: Ihre Recheninfrastrukturen sind begrenzt, veralten rasch und können mit den Kapazitäten der Industrie nicht Schritt halten. Da die industrielle Forschung im Bereich KI zunehmend eine führende Rolle übernimmt, besteht die Gefahr, dass die akademische Forschung ins Hintertreffen gerät. Um ihre Wettbewerbsfähigkeit zu sichern, muss sie daher neue Wege erschliessen. Die bestehenden akademischen Rechenressourcen werden vielfach als unzureichend eingeschätzt. Einige Expertinnen und Experten gehen deshalb Kooperationen mit Industriepartnern ein, um besseren Zugang zu Rechenkapazitäten zu erhalten. Die Initiative «Swiss AI» wird eher als Gewinn für den ETH-Bereich denn als nationale Initiative mit breitem Zugang zu modernsten Rechenressourcen wahrgenommen.

Die Hochschulen betonen, dass neben der Planung der Recheninfrastruktur auch die strategische Planung des Datenzugangs, der Datenspeicherung und der Datenarchivierung von zentraler Bedeutung ist. Bedenken hinsichtlich Datenschutz, Datensicherheit, Souveränität und Kosten führen dabei zu einer Präferenz für on-premise Rechenkapazitäten. Die natürlichen Ressourcen in der Schweiz werden als unzureichend oder zu kostspielig für international wettbewerbsfähiges Hochleistungsrechnen (Supercomputing) wahrgenommen. Angesichts des stark umkämpften Arbeitsmarkts ist es zudem schwierig, erfahrene KI-Fachkräfte zu halten, was zu einer Abwanderung von Wissenschaftlerinnen und Wissenschaftlern in die Industrie geführt hat.

Aus den nationalen und internationalen Interviews geht hervor, dass die vielfältige akademische Landschaft der Schweiz derzeit nicht optimal auf künftige rechenintensive Forschung, wie z. B. KI, vorbereitet zu sein scheint. Die Gestaltung zukunftsorientierter Hochleistungsrechenkapazitäten wird jedoch auf politischer und strategischer Ebene thematisiert.<sup>16,17,18,19,20</sup> Angesichts des raschen Wandels in diesem Bereich und seiner unumstrittenen Bedeutung für die Schweiz empfiehlt der SWR die Entwicklung einer langfristigen KI-Infrastrukturstrategie sowie den Aufbau eines darauf abgestimmten, mehrstufigen nationalen Recheninfrastruktursystems. Dies erfordert die Einrichtung eines nationalen strategischen Leitungsorgans und die Sicherstellung einer angemessenen Finanzierung. Als Sofortmassnahme empfiehlt der SWR, den Aufbau einer Recheninfrastruktur für die akademische Forschung zu einer nationalen Aufgabe zu definieren. Sowohl bei der Ausgestaltung als auch bei der Finanzierung sind neue Mechanismen erforderlich, um die identifizierten Herausforderungen zu bewältigen. Die Infrastruktur soll allen Akteurinnen und Akteuren der BFI-Landschaft gleichberechtigt zugänglich sein und ihren Bedürfnissen Rechnung tragen, wobei zugleich wirtschaftspolitische Strategien berücksichtigt werden müssen. Ein wirtschaftlicher Mehrwert kann beispielsweise durch die Einbindung und Unterstützung von KMU entstehen. Hierfür ist ein strukturierter Dialog zwischen den relevanten Akteurinnen und Akteuren (z. B. Hochschulen, Bund, Kantone) notwendig, um geeignete Mechanismen auf nationaler Ebene zu entwickeln, zu koordinieren, zu implementieren und langfristig zu finanzieren.

## Empfehlungen

### Nationale Strategie – Leitlinien

Die Schweiz muss eine langfristige KI-Infrastrukturstrategie entwickeln, die alle öffentlichen Interessen im BFI-Bereich abdeckt. Diese nationale Strategie muss eine modulare Recheninfrastruktur umfassen, die bestehende Synergien nutzt, indem sie die Bedürfnisse der nationalen grossskaligen Rechenressourcen im BFI-Bereich voraussieht, aufeinander abstimmt und gleichzeitig die Teilnahme der Schweiz an der Entwicklung internationaler Recheninfrastrukturen ermöglicht. Sie soll aufzeigen, wie ein mehrstufiges Recheninfrastruktursystem die Anforderungen der Nutzerinnen und Nutzer an Flexibilität, Skalierbarkeit und Effizienz, einschliesslich Kosteneffizienz, erfüllt und dabei resilient und nachhaltig im Hinblick auf die verfügbaren natürlichen Ressourcen bleibt. Dabei sollen digitale Souveränität und Wissenssicherheit gewahrt sowie Partnerschaften mit werteteilenden Akteurinnen und Akteuren gefördert werden. Zudem soll die Strategie eine Data-Lifecycle-Management-Strategie für das mehrstufige Recheninfrastruktursystem enthalten, die auf den FAIR<sup>21</sup>-Prinzipien aufbaut, und darlegen, wie in der Schweiz eine international konkurrenzfähige Spitzenkompetenz im Bereich KI gefördert und langfristig gesichert werden kann.

- 16 Bundeskanzlei (BK). Strategie Digitale Schweiz 2025. 10. Januar 2025 BBl 2025 31. <https://www.fedlex.admin.ch/eli/fga/2025/31/de> und <https://digital.swiss/de/strategie/fokusthema/kunstliche-intelligenz> (zuletzt geprüft am 26.01.2026).
- 17 Bundesamt für Kommunikation BAKOM. Auslegeordnung zur Regulierung von künstlicher Intelligenz. Bericht an den Bundesrat. 12. Februar 2025. <https://www.bj.admin.ch/bj/de/home/staat/gesetzgebung/kuenstliche-intelligenz.html> (zuletzt geprüft am 26.01.2026).
- 18 Bundeskanzlei (BK). Strategie Einsatz von KI-Systemen in der Bundesverwaltung. 21. März 2025. <https://www.bk.admin.ch/bk/de/home/digitale-transformation-ikt-lenkung/vorgaben/sb021-strategie-einsatz-von-ki-systemen-in-der-bundesverwaltung.html> (zuletzt geprüft am 26.01.2026).
- 19 Eidgenössische Finanzkontrolle (EFK). Internationaler Parallelaudit zur Künstlichen Intelligenz. EFK-25128. 28. Juli 2025. [https://www.efk.admin.ch/wp-content/uploads/publikationen/berichte/wirtschaft\\_und\\_verwaltung/informatikprojekte/25128/25128\\_endgueltige\\_fassung\\_v04.pdf](https://www.efk.admin.ch/wp-content/uploads/publikationen/berichte/wirtschaft_und_verwaltung/informatikprojekte/25128/25128_endgueltige_fassung_v04.pdf) (zuletzt geprüft am 26.01.2026).
- 20 Bundeskanzlei (BK). Weiterentwicklung der Koordination für den Einsatz von künstlicher Intelligenz in der Bundesverwaltung. Konzept. 12. Dezember 2025. <https://www.bk.admin.ch/dam/bk/en/dokumente/dti/themen/ki/konzept-weiterentwicklung-koordination-ki.pdf.download.pdf/konzept-weiterentwicklung-koordination-ki.pdf> (zuletzt geprüft am 26.01.2026).
- 21 FAIR: Findable, accessible, interoperable, reusable. Wilkinson, M., Dumontier, M., Aalbersberg, I. et al. The FAIR Guiding Principles for scientific data management and stewardship. *Sci Data* 3, 160018 (2016). FAIR4RS. Barker, M., Chue Hong, N.P., Katz, D.S. et al. Introducing the FAIR Principles for research software. *Sci Data* 9, 622 (2022).

Für die Erstellung einer solchen nationalen Strategie empfiehlt der SWR, dass die Rechenressourcen alle notwendigen Komponenten, Dienstleistungen und das Personal umfassen, die erforderlich sind, um sie für die Endnutzer nutzbar zu machen. Diese Ressourcen bedürfen daher eines langfristigen Engagements und kontinuierlicher Investitionen. Der SWR ist daher der Ansicht, dass die Nutzung von KI für BFI-Akteurinnen und -Akteure sowie andere rechenintensive akademische Forschung eine geeignete Recheninfrastruktur erfordert, die Synergien nutzt und einer nationalen Strategie folgt. Um dies zu gewährleisten, hält der SWR den Aufbau einer Recheninfrastruktur in Form eines mehrstufigen Systems für unerlässlich. Dies erfordert die Einrichtung eines nationalen strategischen Leitungsorgans sowie die Sicherstellung einer angemessenen Finanzierung.

Der SWR empfiehlt:

## 1. Mehrstufiges Recheninfrastruktursystem

- Aufbau eines mehrstufigen Recheninfrastruktursystems auf der Grundlage einer nationalen Strategie, das von den Beteiligten gemäss den oben genannten Grundsätzen umgesetzt wird.

Stufen 0 und 1<sup>22</sup>: Internationales und nationales Supercomputing<sup>23</sup>

- Weiterentwicklung der bestehenden nationalen und internationalen Supercomputing-Strukturen durch Erweiterung des Zugangs zu und der Beteiligung an den international leistungsfähigsten Rechenressourcen und deren Integration mit den nationalen Ressourcen.
- Den bestehenden nationalen Supercomputing-Anbieter CSCS, der zugleich den Zugang zu internationalen Supercomputing-Einrichtungen<sup>24</sup> bietet, einer Governance-Struktur zu unterstellen, die seinem nationalen Auftrag entspricht, um die Versorgung aller BFI-Akteure in der gesamten Schweiz sicherzustellen.

Stufe 2: Nationale Hochleistungsrecheninfrastruktur

- Förderung einer nationalen Hochleistungsrecheninfrastruktur, die einer breiten Nutzergruppe zur Verfügung steht, deren Anforderungen keine Supercomputer-Ressourcen erfordern, und die bestehende Rechenressourcen integriert oder auf diesen aufbaut.

Stufe 3: Regionale Rechenzentren<sup>25</sup>

- Einrichtung lokaler oder regionaler KI-Expertengruppen an Hochschuleinrichtungen, die die Nutzerinnen und Nutzer vor Ort unterstützen und sie bei Bedarf an die entsprechenden Rechenressourcen innerhalb des mehrstufigen Systems verweisen.
- Die Stufen stellen Rechenressourcen bereit, die sämtliche notwendigen Komponenten, Dienste und Ressourcen umfassen, die erforderlich sind, um die Rechenleistung für die Endnutzer nutzbar zu machen. Zugleich fungieren sie als nationaler Pool von Expertinnen und Experten zur Unterstützung der Nutzenden und der Wissenschaft, einschliesslich im Bereich der KI-Entwicklung.

Damit wird der vielfältigen Schweizer Hochschullandschaft eine umfassende, angemessen skalierte, nachhaltige und belastbare Recheninfrastruktur bereitgestellt, die eine breite Palette von Anforderungen abdeckt. Der SWR stellt sich ein mehrstufiges Recheninfrastruktursystem vor, das durch Interoperabilität eine höhere Effizienz ermöglicht und den Nutzenden einen nahtlosen Übergang zwischen den Stufen bietet (z. B. für föderales Lernen). Das System soll – unter Berücksichtigung der unterschiedlichen Finanzierungsquellen und bestehenden Governance-Strukturen – über einen einheitlichen Zugangspunkt verwaltet werden (z. B. für die Stufen 0, 1 und 2).

22 Beispiel: Nationales Hochleistungsrechenzentrum (CSCS) mit Verbindung zur internationalen Supercomputing-Einrichtung LUMI (FIN); nationales Supercomputing: Alps.

23 In diesem Kontext bezeichnet «Supercomputing» das höchstleistungsfähige Rechnen, also die fortschrittlichsten Systeme, die an den äussersten Grenzen der heutigen Technologie operieren.

24 Zugang zu LUMI: <https://www.lumi-supercomputer.eu/get-started-2021/users-in-switzerland/> und durch die Aktivitäten von HELvetic AI Resources, Technologies and Services (HEARTS), siehe <https://csc.fi/en/news/eurohpc-j-selects-ai-factory-antennas-to-broaden-ai-factories-initiative/> und <https://goba.swiss/en/switzerland-joins-european-ai-factory-network-with-new-hearts-antenna/> (zuletzt geprüft am 26.01.26).

25 Beispiel: Hochschulen: zentrale HPC-Infrastrukturen (On-Premise-Lösungen) der kantonalen Universitäten oder der Fachhochschulen.

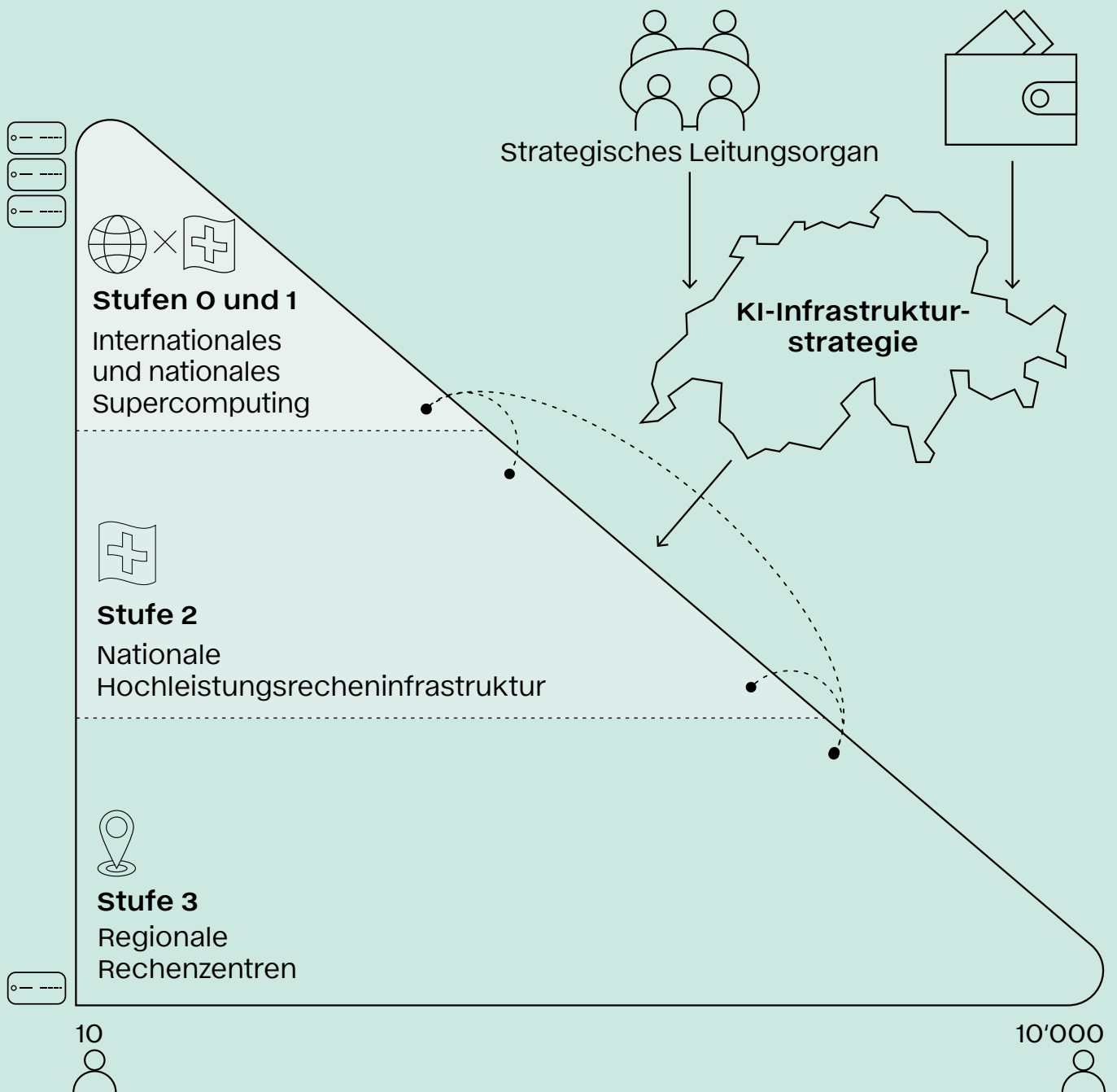


Abbildung 2: Illustration des Systems der mehrstufigen Recheninfrastruktur. Die X-Achse veranschaulicht die Anzahl der Nutzerinnen und Nutzer des Systems, während die Y-Achse den Grad der Rechenkapazität angibt. Die auf der X-Achse dargestellte Nutzerzahl entspricht nicht den tatsächlichen Zahlen, sondern dient lediglich der Veranschaulichung. Die gestrichelte Linie zeigt die Durchlässigkeit zwischen den einzelnen Stufen. Die gestrichelten Halbkreise stehen für die Interoperabilität zwischen den Stufen. Interoperabilität bezieht sich hier auf die Fähigkeit, Informationen über alle Stufen hinweg auszutauschen und zu nutzen, z. B. Daten (einschliesslich Software) und Rechenkapazität.

## 2. Strategisches Leitungsorgan für eine agile Recheninfrastruktur

- Einrichtung eines nationalen, unabhängigen und neutralen strategischen Leitungsorgans, das die Ausgestaltung einer zukunftsorientierten Recheninfrastruktur in Form eines mehrstufigen Systems bewertet und darüber entscheidet. Das strategische Leitungsorgan ist zuständig für:
  - die strategische Gestaltung des mehrstufigen Systems auf der Grundlage einer kontinuierlichen Bewertung des nationalen Bedarfs an grossskaligen Rechenressourcen sowie der internationalen Entwicklungen im Bereich Recheninfrastrukturen unter Berücksichtigung der Multi-Level-Governance. Die nationale Strategie sollte auf den oben dargelegten Leitprinzipien basieren;
  - die Überwachung der Umsetzung und Leistungsfähigkeit der nationalen Strategie durch die einzelnen Stufen des mehrstufigen Recheninfrastruktursystems.

Der Bundesrat sollte dieses strategische Leitungsorgan als unabhängige Expertenkommission einsetzen. Es sollte mit den notwendigen Befugnissen ausgestattet werden, um seine Aufgaben zu erfüllen, gegebenenfalls auf gesetzlicher Grundlage.

- Das strategische Leitungsorgan sollte sämtliche für die Wahrnehmung seiner Aufgaben erforderlichen Kompetenzen bündeln. Dazu zählen insbesondere technisches<sup>26</sup>, wissenschaftliches<sup>27</sup>, strategisches<sup>28</sup>, juristisches<sup>29</sup> und finanzielles<sup>30</sup> Fachwissen. Gleichzeitig sollte es die Nutzenden sowie die relevanten Interessengruppen der nationalen und regionalen Rechenressourcen der Hochschulen angemessen vertreten. Die Grösse des Leitungsorgans ist so zu bemessen, dass eine effiziente Steuerung gewährleistet ist.

26 Technik: fortgeschrittene Hardware-Architekturen, Beschleuniger und Spezialchips, Systemsoftware und Programmierung, Optimierung und Leistungsabstimmung, Netz- und Dateninfrastrukturen, Datenmanagement und neue Technologien (z. B. Quantencomputer).

27 Wissenschaftlich: Domänenspezifisches Forschungs-Know-how: Verständnis der einzigartigen Rechen-, Daten- und Simulationsanforderungen verschiedener Forschungsbereiche, um sicherzustellen, dass HPC-Systemdesigns mit realen wissenschaftlichen Anforderungen übereinstimmen.

28 Strategisch: Interdisziplinäre Zusammenarbeit und strategischer Weitblick: Förderung der Zusammenarbeit zwischen Expertinnen und Experten aus Technik, Wissenschaft, Politik und Wirtschaft bei gleichzeitiger Wahrung einer langfristigen, strategischen Vision, um die Computerinfrastruktur an die sich entwickelnden Forschungs- und Markttrends anzupassen.

29 Rechtliche Aspekte: Dynamik des Ökosystems und der Lieferkette: Verständnis der Lieferantenbeziehungen, Überlegungen zum geistigen Eigentum und Einhaltung der rechtlichen Rahmenbedingungen, um Risiken zu mindern und eine rechtlich einwandfreie Recheninfrastruktur zu gewährleisten, einschliesslich der zugehörigen Speicher- und Archivierungssysteme. Risikomanagement und Einhaltung von Vorschriften: Management von Rechts- und Cybersicherheitsrisiken, einschliesslich der Einhaltung von Exportkontrollen, Datenschutzbestimmungen und anderen rechtlichen Verpflichtungen, um sicherzustellen, dass strategische Investitionen sowohl widerstandsfähig als auch gesetzeskonform sind.

30 Finanzen: Strategische Investitionsanalyse und finanzieller Sachverstand: Fähigkeit zur Durchführung gründlicher Kosten-Nutzen-Analysen und zur Verwaltung langfristiger Investitionsstrategien unter Berücksichtigung von Leistung, Skalierbarkeit und finanzieller Verantwortung.

### 3. Finanzielle Implikationen

- Für eine Finanzierung, die eine vollständige Kostendeckung der Rechenressourcen ermöglicht, können projektbezogene Fördermittel nur ergänzend herangezogen werden, da sie in der Regel kurzfristig angelegt sind. Rechenressourcen erfordern jedoch ein langfristiges Engagement und nachhaltige Investitionen. Daher ist für alle Stufen der Recheninfrastruktur eine verlässliche, langfristige Finanzierungsgrundlage sicherzustellen. Grundsätzlich sollten Rechenressourcen transparent und kosteneffizient bereitgestellt werden. Dies bedeutet, dass die entstehenden Kosten vollständig gedeckt werden müssen, aber keine Gewinne aus Hochschuleinrichtungen erzielt werden sollten.

# Résumé

Au cours de la dernière décennie, la recherche fondée sur les données a gagné en importance, notamment grâce à des méthodes d'intelligence artificielle (IA) telles que l'apprentissage automatique. Une infrastructure de calcul adéquate est essentielle pour mener des analyses exigeant d'importantes données et ressources en calcul. Le Conseil suisse de la science (CSS) a vérifié si le monde académique en Suisse disposait des ressources nécessaires pour satisfaire ses besoins actuels et futurs. Par le biais des entretiens qu'il a menés auprès d'organisations en Suisse et à l'étranger, il a recueilli des informations sur les infrastructures de calcul et services cloud utilisés actuellement ainsi que sur les perspectives et projets futurs. Les hautes écoles suisses réfléchissent activement à la meilleure façon d'adapter leurs systèmes aux exigences croissantes en matière de capacité de calcul. Plusieurs défis ont été identifiés autour de cette question d'importance stratégique pour l'avenir du système de formation, de recherche et d'innovation (FRI) de la Suisse.

La planification des ressources est une tâche complexe car les futurs besoins en matière de calcul et les effets de l'optimisation technique sont difficiles à prévoir. Les hautes écoles sont confrontées à des défis dans la mesure où leur parc d'infrastructures de calcul est limité, leurs infrastructures deviennent rapidement obsolètes et sont moins compétitives que celles des entreprises. L'industrie étant de plus en plus à la pointe de la recherche en IA, le monde académique risque de prendre du retard et doit donc trouver des moyens de rester compétitif. De plus, il a la réputation de disposer de ressources en calcul insuffisantes, ce qui amène certains spécialistes à collaborer avec des entreprises afin de s'assurer un meilleur accès aux ressources. L'initiative suisse en matière d'IA est perçue comme une initiative profitant principalement au domaine des écoles polytechniques fédérales (EPF), plutôt qu'une initiative nationale misant sur les ressources en calcul les plus avancées.

Les hautes écoles soulignent combien il est important de planifier l'accès, le stockage et l'archivage des données parallèlement à la planification de l'infrastructure de calcul. Les préoccupations quant à la confidentialité et à la protection des données, à la souveraineté et au coût incitent à préférer un traitement des données en interne. Les ressources naturelles de la Suisse sont considérées comme insuffisantes ou trop coûteuses pour le calcul à haute performance dont il est habituellement question à l'échelle internationale. Il est difficile de retenir les professionnels de l'IA sur un marché du travail concurrentiel ; ceci se traduit par la fuite des cerveaux du monde académique vers l'industrie.

Au regard des résultats des entretiens menés en Suisse et à l'étranger, le paysage académique de la Suisse, très diversifié, semble actuellement mal préparé à de futures activités de recherche nécessitant d'importantes ressources en calcul telles que l'IA. Cela dit, la conception de capacités de calcul à haute performance à même de répondre aux besoins de demain est traitée au niveau politique et stratégique<sup>31,32,33,34,35</sup>. Étant donné que ce domaine évolue rapidement et que son importance pour la Suisse est indéniable, le CSS recommande d'élaborer une stratégie à long terme en matière d'infrastructure d'IA et de créer un système d'infrastructure de calcul à plusieurs niveaux sur la base de cette stratégie nationale. Pour ce faire, il est nécessaire de mettre en place un comité stratégique national et de garantir un financement adéquat. À titre de mesure immédiate, le CSS recommande la mise sur pied d'une infrastructure de calcul pour la recherche académique comme tâche nationale, considérant la création de nouveaux mécanismes, à la fois en termes de conception et de financement, afin de relever les défis qui ont été identifiés. Il s'agira de rendre cette infrastructure accessible sans distinction aux différents acteurs FRI en fonction de leurs besoins, tout en tenant compte des stratégies économiques, de sorte que cette infrastructure profite aussi à l'économie, par exemple par l'intégration et l'accompagnement de petites et moyennes entreprises (PME). À cette fin, les acteurs concernés, tels que les hautes écoles, la Confédération ou les cantons, doivent engager un dialogue afin de développer, de coordonner, d'établir et de financer les mécanismes nécessaires à l'échelle nationale sur le long terme.

- 31 Chancellerie fédérale ChF. Stratégie Suisse numérique 2025. 10 janvier 2025. FF 2025 31. <https://www.fedlex.admin.ch/eli/fga/2025/31/fr> and <https://digital.swiss/fr/strategie/themes-prioritaires.html#intelligence-artificielle> (consulté le 26 janvier 2026).
- 32 Office fédéral de la communication OFCOM. État des lieux sur la réglementation de l'intelligence artificielle. Rapport à l'attention du Conseil fédéral. 12 février 2025. <https://www.bj.admin.ch/bj/fr/home/staat/gesetzgebung/kuenstliche-intelligenz.html> (consulté le 26 janvier 2026).
- 33 Chancellerie fédérale ChF. Stratégie pour l'utilisation de systèmes d'IA dans l'administration fédérale. 21 mars 2025. <https://www.bk.admin.ch/bk/fr/home/digitale-transformation-ikt-lenkung/vorgaben/sb021-strategie-einsatz-von-ki-systemen-in-der-bundesverwaltung.html> (consulté le 26 janvier 2026).
- 34 Contrôle fédéral des finances CDF. Audit parallèle international sur l'intelligence artificielle. CDF-25128. 28 juillet 2025. [https://www.efk.admin.ch/wp-content/uploads/publikationen/berichte/wirtschaft\\_und\\_verwaltung/informatikprojekte/25128/25128\\_endgueltige\\_fassung\\_v04.pdf](https://www.efk.admin.ch/wp-content/uploads/publikationen/berichte/wirtschaft_und_verwaltung/informatikprojekte/25128/25128_endgueltige_fassung_v04.pdf) (rapport en allemand) ; [https://www.efk.admin.ch/wp-content/uploads/publikationen/berichte/wirtschaft\\_und\\_verwaltung/informatikprojekte/25128/25128\\_wik-f.pdf](https://www.efk.admin.ch/wp-content/uploads/publikationen/berichte/wirtschaft_und_verwaltung/informatikprojekte/25128/25128_wik-f.pdf) (résumé en français) (consulté le 26 janvier 2026).
- 35 Chancellerie fédérale ChF. Renforcement de la coordination en matière d'utilisation de l'intelligence artificielle dans l'administration fédérale. Concept. 12 décembre 2025. [https://www.bk.admin.ch/bk/fr/home/digitale-transformation-ikt-lenkung/kuenstliche\\_intelligenz.html](https://www.bk.admin.ch/bk/fr/home/digitale-transformation-ikt-lenkung/kuenstliche_intelligenz.html) (consulté le 26 janvier 2026).

## Recommandations

### Stratégie nationale – principes directeurs

La Suisse a besoin d'une stratégie à long terme en matière d'infrastructure d'IA qui réponde à tous les besoins publics dans le domaine FRI. Cette stratégie nationale devra inclure une infrastructure de calcul modulaire utilisant les synergies existantes. Elle doit anticiper et harmoniser les besoins des grandes infrastructures de calcul nationales du domaine FRI et permettre à la Suisse de participer au développement d'infrastructures de calcul internationales. Il s'agit de décrire comment un système d'infrastructure de calcul à plusieurs niveaux peut satisfaire aux exigences des utilisateurs en matière de flexibilité, d'évolutivité et d'efficacité, incluant l'efficacité des coûts, tout en étant résilient et durable au regard des ressources naturelles. Cette stratégie doit valoriser la souveraineté numérique ainsi que la sécurité des connaissances, et favoriser les partenariats avec les acteurs participant à la création de valeur. Elle doit également inclure une stratégie en matière de gestion du cycle de vie des données pour le système d'infrastructure de calcul à plusieurs niveaux, stratégie qui repose sur les principes FAIR<sup>36</sup>, et expliquer comment encourager et maintenir en Suisse une expertise de premier ordre en matière d'IA, à même de faire face à la concurrence sur le plan international.

Afin d'élaborer une telle stratégie nationale, le CSS recommande d'inclure dans les ressources en calcul tous les moyens qui doivent être mobilisés (matériel, services, personnel) de manière à les mettre à disposition des utilisateurs finaux. Par conséquent, cette approche suppose un engagement et des investissements à long terme. C'est pourquoi le CSS estime que l'utilisation de l'IA par les acteurs du domaine FRI, et dans le cadre d'autres activités de recherche académiques exigeant d'importantes ressources en calcul, implique de disposer d'une infrastructure de calcul appropriée qui exploite les synergies et relève d'une stratégie nationale. Il considère ainsi comme essentiel de créer une infrastructure de calcul sous la forme d'un système à plusieurs niveaux. Pour ce faire, il est nécessaire de mettre en place un comité stratégique national et de garantir un financement adéquat.

<sup>36</sup> FAIR : Findable, Accessible, Interoperable, Reusable (Facilement trouvable, Accessible, Interopérable, Réutilisable). Wilkinson, M., Dumontier, M., Aalbersberg, I. et al. The FAIR Guiding Principles for scientific data management and stewardship. *Sci Data* 3, 160018 (2016). FAIR4RS. Barker, M., Chue Hong, N.P., Katz, D.S. et al. Introducing the FAIR Principles for research software. *Sci Data* 9, 622 (2022).

Le CSS formule les recommandations ci-après.

## 1. Système d'infrastructure de calcul à plusieurs niveaux

- Mettre en place un système d'infrastructure de calcul à plusieurs niveaux, fondé sur une stratégie nationale, mis en œuvre par les acteurs concernés et conforme aux principes énoncés ci-dessus.

Pour les niveaux 0 et 1<sup>37</sup> : Supercalculateurs nationaux et internationaux<sup>38</sup>

- Continuer à développer les niveaux actuels des supercalculateurs nationaux et internationaux en élargissant l'accès et la participation aux ressources en calcul internationales les plus sophistiquées, et en les intégrant aux ressources nationales.
- Placer le fournisseur national en matière de supercalculateur (CSCS), qui donne également accès aux supercalculateurs internationaux<sup>39</sup>, sous une gouvernance garante de son mandat national de prestations à fournir aux acteurs du domaine FRI de toute la Suisse.

Pour le niveau 2 : Ordinateur national à haute performance

- Encourager une infrastructure nationale de calcul à haute performance au service d'une grande variété d'utilisateurs qui peuvent se passer de supercalculateurs et intègrent ou développent les ressources informatiques existantes.

Pour le niveau 3 : Centres de calcul régionaux<sup>40</sup>

- Mettre en place des groupes locaux ou régionaux d'experts en IA dans les hautes écoles afin d'aider directement les utilisateurs et de les orienter, le cas échéant, vers les ressources informatiques appropriées au sein du système à plusieurs niveaux.
- Les niveaux devraient fournir des ressources en calcul englobant tous les composants, services et ressources nécessaires pour mettre à la disposition des utilisateurs finaux une puissance de calcul exploitable et, dans le même temps, fonctionner comme des pools d'experts nationaux pour le soutien aux utilisateurs et aux chercheurs, y compris pour les développements de l'IA.

Le paysage académique suisse, très diversifié, disposera ainsi d'une infrastructure de calcul complète, correctement dimensionnée, durable, résiliente et en mesure de répondre à un large éventail de besoins. Le CSS envisage un système d'infrastructure de calcul à plusieurs niveaux qui permette une plus grande efficacité en s'appuyant sur l'interopérabilité et offrant aux utilisateurs la possibilité de passer facilement d'un niveau à l'autre, par exemple pour l'apprentissage fédéré. Ce système est géré par un point d'accès unique (notamment pour les niveaux 0, 1 et 2), tout en tenant compte des différentes sources de financement et des structures de gouvernance déjà en place.

37 Exemple : CSCS : lien avec les supercalculateurs internationaux – LUMI (FIN) ; supercalculateur suisse : Alps.

38 Dans ce contexte, le terme supercalculateur se réfère aux calculateurs les plus perfectionnés, c'est-à-dire aux systèmes les plus avancés fonctionnant aux frontières extrêmes des technologies actuelles.

39 Accès à LUMI : <https://www.lumi-supercomputer.eu/get-started-2021/users-in-switzerland/> et au travers des activités de Helvetic AI Resources, Technologies and Services (HEARTS), voir <https://csc.fi/en/news/eurohpc-ju-se-lects-ai-factory-antennas-to-broaden-ai-factories-initiative/> et <https://goba.swiss/en/switzerland-joins-european-ai-factory-network-with-new-hearts-antenna/> (consulté le 26 janvier 2026).

40 Exemple : Hautes écoles : systèmes HPC centraux (solutions sur site) des universités cantonales ou des hautes écoles spécialisées.

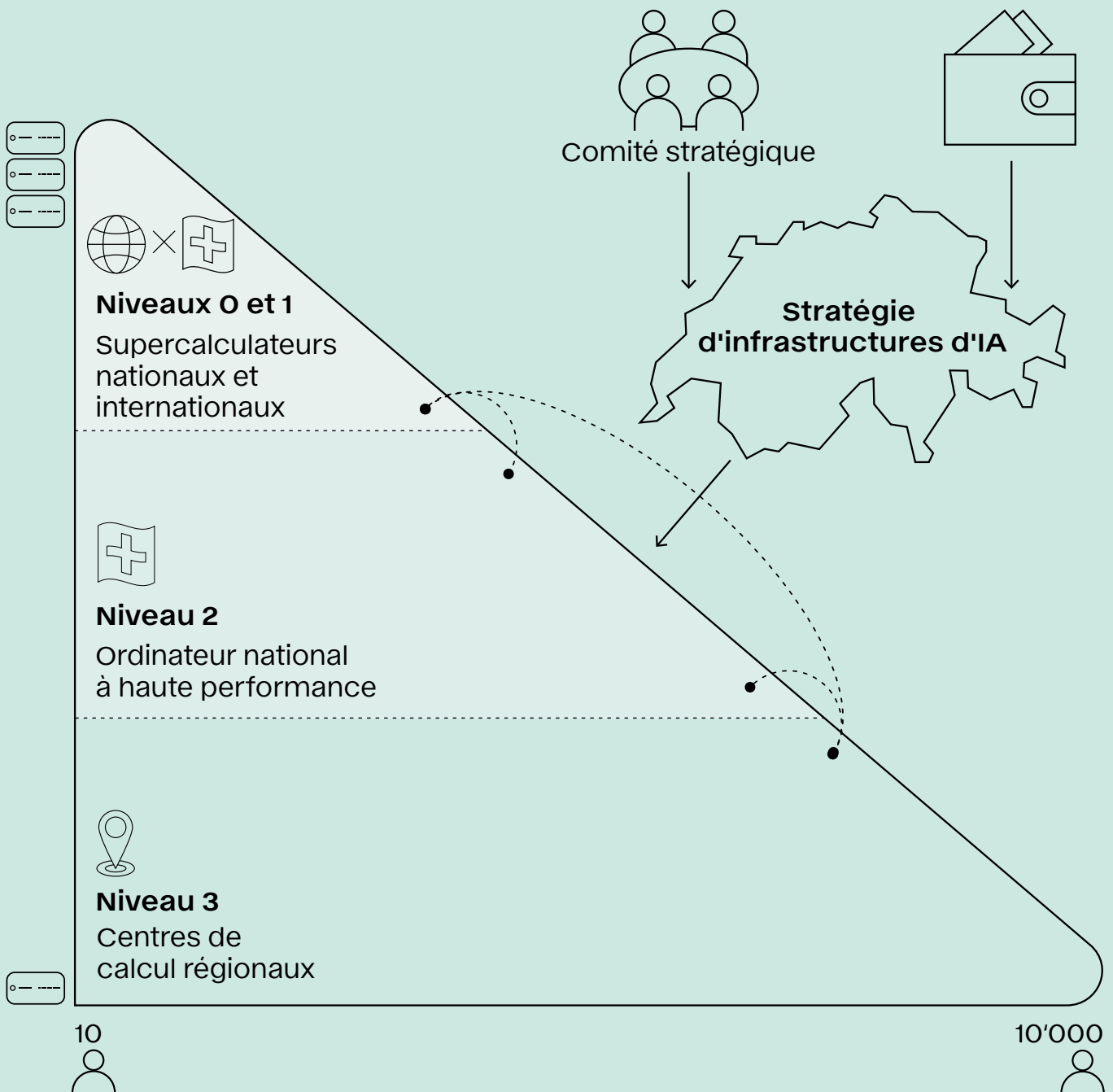


Figure 3 : Représentation schématisque du système d'infrastructure de calcul à plusieurs niveaux.

L'axe horizontal correspond au nombre d'utilisateurs du système et l'axe vertical au degré de la capacité de calcul. Le nombre d'utilisateurs indiqué sur l'axe horizontal ne reflète pas le nombre réel d'utilisateurs. Il a seulement valeur d'exemple. Les lignes en pointillé illustrent la perméabilité entre les différents niveaux. Les demi-cercles en pointillés illustrent l'interopérabilité. La notion d'interopérabilité fait ici référence à la possibilité d'échanger et d'utiliser des informations à tous les niveaux, par exemple concernant la capacité de données (y compris les logiciels) et la capacité de calcul.

## 2. Comité stratégique pour une infrastructure de calcul agile

- Créer un comité stratégique national, indépendant et neutre, chargé de définir la conception d'une infrastructure de calcul adaptée aux besoins à venir sous la forme d'un système à plusieurs niveaux. Ce comité stratégique aurait les deux missions décrites ci-après.
  - Concevoir le système à plusieurs niveaux de manière stratégique, en se basant sur l'évaluation en continu des besoins en grandes infrastructures de calcul nationales et des évolutions des infrastructures de calcul internationales, tout en tenant compte de la gouvernance multiniveaux. La stratégie nationale devrait reposer sur les principes directeurs exposés précédemment.
  - Surveiller la mise en œuvre et l'efficacité de la stratégie nationale à tous les niveaux du système d'infrastructure de calcul.

Le Conseil fédéral devrait nommer ce comité stratégique en tant que commission d'experts indépendante. Ce comité devrait être doté des pouvoirs nécessaires pour s'acquitter de ses tâches, le cas échéant selon une base légale.

- Le comité stratégique doit regrouper toutes les compétences dont il a besoin pour accomplir ses tâches, notamment l'expertise technique<sup>41</sup>, scientifique<sup>42</sup>, stratégique<sup>43</sup>, juridique<sup>44</sup> et financière<sup>45</sup>, représenter les utilisateurs et les groupes d'intérêts des infrastructures de calcul nationales et régionales des hautes écoles et avoir une taille adéquate pour garantir un pilotage efficace.

## 3. Incidences financières

- Adapter le mécanisme financier de façon à garantir une prise en charge de l'intégralité des ressources en calcul. Le financement ne peut s'appuyer que partiellement sur des subventions car celles-ci couvrent uniquement les coûts à court terme. Les ressources en calcul nécessitent en effet un engagement et des investissements à long terme. Par conséquent, un financement à long terme adéquat doit être assuré pour tous les niveaux. En général, les ressources en calcul doivent être mises à la disposition des utilisateurs de manière transparente et efficace, ce qui signifie que les coûts doivent être couverts, mais que les hautes écoles ne doivent pas faire de bénéfices.

41 Expertise technique : architectures matérielles avancées, accélérateurs et puces spéciales, logiciels et programmes, optimisation et réglage des performances, infrastructures de réseaux et de données, gestion des données et technologies émergentes (p. ex. informatique quantique).

42 Expertise scientifique : expertise dans un domaine de recherche spécifique : être à même de comprendre les besoins particuliers en matière de calcul, de données et de simulation des différents domaines de recherche afin de s'assurer que la conception du système HPC répond aux attentes réelles du monde scientifique.

43 Expertise stratégique : collaboration interdisciplinaire et vision stratégique : être à même de favoriser la collaboration entre les experts techniques, scientifiques, politiques et économiques, tout en maintenant une vision stratégique à long terme afin d'adapter l'infrastructure de calcul aux évolutions de la recherche et aux tendances du marché.

44 Expertise juridique : dynamique de l'écosystème et de la chaîne d'approvisionnement : être à même de comprendre les relations avec les fournisseurs et de traiter les questions de propriété intellectuelle et de respect des cadres réglementaires afin d'atténuer les risques et de garantir une infrastructure de calcul juridiquement saine, y compris les systèmes de stockage et d'archivage associés. Gestion des risques et conformité réglementaire : être à même de gérer les risques juridiques et de cybersécurité, y compris le respect des contrôles à l'exportation, des réglementations en matière de protection des données et d'autres obligations légales, afin de garantir que les investissements stratégiques sont à la fois résilients et conformes.

45 Expertise financière : analyse stratégique des investissements et habileté financière : être à même de mener des analyses coûts/bénéfices approfondies et de gérer des stratégies d'investissement à long terme en tenant compte des résultats, de l'évolutivité et de la responsabilité fiscale.

# Riassunto

Nell'ultimo decennio l'importanza della ricerca basata sui dati è cresciuta, grazie anche a metodi di intelligenza artificiale (IA) come l'apprendimento automatico. Un'adeguata infrastruttura di calcolo è indispensabile per analisi che richiedono grandi quantità di dati e una notevole potenza computazionale. Il Consiglio svizzero della scienza (CSS) ha condotto un'indagine per valutare se il mondo accademico svizzero disponga delle risorse necessarie per soddisfare i suoi fabbisogni attuali e futuri. Attraverso interviste con organizzazioni in Svizzera e all'estero, ha raccolto informazioni sulle attuali infrastrutture di calcolo, sui servizi cloud e sulle visioni e i piani futuri. Le scuole universitarie svizzere stanno valutando concretamente come adeguare i propri sistemi alle crescenti esigenze computazionali. Questo adeguamento è considerato determinante sul piano strategico per il futuro del settore ERI (educazione, ricerca e innovazione) in Svizzera e sono state identificate diverse sfide.

Non è semplice prevedere quali risorse saranno necessarie, perché è difficile stimare le future esigenze di calcolo e gli effetti dell'ottimizzazione. Le scuole universitarie devono affrontare diverse sfide dovute al fatto che le loro infrastrutture di calcolo sono limitate, diventano rapidamente obsolete e sono meno competitive rispetto a quelle dell'industria. Dato che quest'ultima assume un ruolo sempre più dominante nella ricerca sull'IA, il mondo accademico deve trovare il modo di rimanere competitivo se non vuole rischiare di rimanere indietro. Le risorse di calcolo accademiche disponibili sono spesso considerate inadeguate, al punto da spingere alcuni esperti a collaborare con il settore privato per ottenere un migliore accesso a capacità di calcolo. L'Iniziativa svizzera per l'IA (Swiss AI Initiative) viene percepita come un intervento a favore del settore dei Politecnici federali (PF), piuttosto che come un progetto d'interesse nazionale associato alle risorse di calcolo più avanzate.

Le scuole universitarie sottolineano l'importanza di pianificare l'accesso ai dati, la loro memorizzazione e archiviazione parallelamente alla progettazione dell'infrastruttura di calcolo. Le preoccupazioni relative alla privacy, alla protezione, alla sovranità e ai costi dei dati inducono a privilegiare soluzioni di calcolo interne. Le risorse naturali svizzere sono considerate insufficienti o troppo costose per il tipo di supercalcolo su larga scala diffuso a livello internazionale. Trattenere professionisti dell'IA è difficile in un mercato del lavoro competitivo, il che porta a una fuga di cervelli dal mondo accademico verso il settore industriale.

Sulla base dei risultati emersi dalle interviste condotte a livello nazionale e internazionale, il panorama accademico svizzero, nella sua diversità, sembra scarsamente preparato per future attività di ricerca che richiedono un uso intensivo delle risorse di calcolo, come quelle sull'IA. Tuttavia, la progettazione di capacità di calcolo ad alte prestazioni in grado di rispondere alle esigenze future viene esaminata a livello politico e strategico<sup>46,47,48,49,50</sup>. Poiché questo settore è soggetto a rapidi cambiamenti e la sua importanza per la Svizzera è indubbia, il Consiglio raccomanda lo sviluppo di una strategia a lungo termine per un'infrastruttura IA e la creazione di un sistema di infrastrutture di calcolo a più livelli basato su questa strategia nazionale. Ciò richiede l'istituzione di un comitato strategico nazionale e la garanzia di finanziamenti adeguati. A tal fine, il Consiglio raccomanda le seguenti misure immediate: dichiarare lo sviluppo di un'infrastruttura di calcolo per la ricerca accademica un compito nazionale per il quale occorre trovare nuovi meccanismi, in termini sia di progettazione che di finanziamento, al fine di far fronte alle sfide individuate; rendere tale infrastruttura accessibile senza distinzione ai diversi attori del settore ERI, in linea con le loro esigenze e compatibile con le strategie economiche, in modo che vada a beneficio anche dell'economia, ad esempio attraverso l'integrazione e l'accompagnamento delle PMI. A questo proposito, i principali portatori di interesse (p. es. scuole universitarie, Confederazione, Cantoni) devono avviare un dialogo per sviluppare, coordinare, istituire e finanziare i meccanismi necessari a livello nazionale con un impegno a lungo termine.

## Raccomandazioni

### Strategia nazionale: principi guida

La Svizzera deve sviluppare una strategia a lungo termine per un'infrastruttura IA che soddisfi tutte le esigenze pubbliche nel settore ERI. Questa strategia nazionale deve includere un'infrastruttura di calcolo modulare che sfrutti le sinergie esistenti, anticipando e allineando le esigenze in materia di risorse di calcolo nazionali su larga scala del settore ERI e consentendo alla Svizzera di partecipare allo sviluppo di infrastrutture di calcolo internazionali. È necessario definire il modo in cui un sistema di infrastrutture di calcolo a più livelli possa soddisfare i requisiti degli utenti in termini di flessibilità, scalabilità ed efficienza, tenendo conto del rapporto costi-benefici, e allo stesso tempo, garantire la resilienza e un uso sostenibile delle risorse naturali. Tale strategia deve promuovere la sovranità digitale e la sicurezza delle conoscenze, favorendo partenariati con portatori di interesse che condividono gli stessi valori. Inoltre, deve includere un piano di gestione del ciclo di vita dei dati per il sistema di calcolo a più livelli fondato sui principi FAIR<sup>51</sup> e deve definire come promuovere e supportare un'eccellenza svizzera nel settore dell'IA competitiva a livello internazionale.

46 Cancelleria federale (CaF), Strategia Svizzera digitale 2025, 10 gennaio 2025. FF 2025 31, <https://www.fedlex.admin.ch/eli/fga/2025/31/it> e <https://digital.swiss/it/strategia/fokusthema.htmlintelligenza-artificiale> (consultato il 26 gennaio 2026).

47 Ufficio federale delle comunicazioni (UFCOM), Regolamentazione dell'intelligenza artificiale: analisi della situazione attuale. Rapporto all'attenzione del Consiglio federale, 12 febbraio 2025, <https://www.bj.admin.ch/bj/it/home/staat/gesetzgebung/kuenstliche-intelligenz.html> (consultato il 26 gennaio 2026).

48 Cancelleria federale (CaF), Strategia per l'utilizzo dei sistemi di IA nell'Amministrazione federale, 21 marzo 2025, <https://www.bk.admin.ch/bk/it/home/digitale-transformation-ikt-lenkung/vorgaben/sb021-strategie-einsatz-von-ki-systemen-in-der-bundesverwaltung.html> (consultato il 26 gennaio 2026).

49 Controllo federale delle finanze (CDF), Audit parallelo a livello internazionale sull'intelligenza artificiale, CDF-25128, 28 luglio 2025, [Audit parallelo a livello internazionale sull'intelligenza artificiale - Controllo federale delle finanze \(CDF\)](#) (rapporto integrale non disponibile in italiano; consultato il 26 gennaio 2026).

50 Cancelleria federale (CaF), Concetto per l'ulteriore sviluppo del coordinamento per l'utilizzo dell'IA nell'Amministrazione federale, 12 dicembre 2025, [https://www.bk.admin.ch/bk/it/home/digitale-transformation-ikt-lenkung/kuenstliche\\_intelligenz.html](https://www.bk.admin.ch/bk/it/home/digitale-transformation-ikt-lenkung/kuenstliche_intelligenz.html) (consultato il 26 gennaio 2026).

51 FAIR: Findable, accessible, interoperable, reusable (trovabili, accessibili, interoperabili, riutilizzabili). Wilkinson, M., Dumontier, M., Aalbersberg, I. et al. The FAIR Guiding Principles for scientific data management and stewardship, *Sci Data* 3, 160018 (2016). FAIR4RS. Barker, M., Chue Hong, N.P., Katz, D.S. et al. Introducing the FAIR Principles for research software, *Sci Data* 9, 622 (2022).

Ai fini dell'elaborazione di tale strategia nazionale, il CSS raccomanda che le risorse di calcolo comprendano tutti i componenti, i servizi e il personale necessari per renderle fruibili agli utenti finali. Di conseguenza, queste risorse richiedono un impegno e investimenti a lungo termine. Per questo motivo il Consiglio ritiene che, per consentire l'uso dell'IA ai portatori di interesse del settore ERI e per altre ricerche accademiche che richiedono una notevole potenza computazionale, sia necessaria un'infrastruttura di calcolo adeguata che sfrutti le sinergie esistenti e segua una strategia nazionale. A tal fine, il Consiglio ritiene che sia fondamentale creare un'infrastruttura sotto forma di un sistema a più livelli. Per sviluppare un sistema di questo tipo è necessario istituire un comitato strategico nazionale e garantire finanziamenti adeguati.

Di seguito sono esposte le raccomandazioni del Consiglio.

## 1. Sistema di infrastrutture di calcolo a più livelli

- Creare un sistema a più livelli basato su una strategia nazionale, realizzato dai portatori di interesse e guidato dai principi sopra descritti.

Per i livelli 0 e 1<sup>52</sup>: supercalcolo internazionale e nazionale<sup>53</sup>

- Sviluppare ulteriormente i livelli esistenti di supercalcolo internazionale e nazionale, ampliando l'accesso e la partecipazione alle risorse di calcolo internazionali più avanzate e integrandole con le risorse nazionali
- Porre l'attuale fornitore nazionale di supercalcolo (CSCS), che fornisce anche l'accesso a strutture internazionali di supercalcolo<sup>54</sup>, sotto una governance che adempia al suo mandato nazionale, al fine di garantire il servizio a tutti i portatori di interesse del settore ERI in tutta la Svizzera

Per il livello 2: calcolo ad alte prestazioni nazionale

- Promuovere un'infrastruttura nazionale di calcolo ad alte prestazioni, accessibile a una vasta gamma di utenti che non richiedono risorse di supercalcolo, e che integri le risorse di calcolo esistenti o si fondi su di esse

Per il livello 3: centri di calcolo regionali<sup>55</sup>

- Creare gruppi di esperti di IA a livello locale e regionale presso scuole universitarie, affinché offrano supporto agli utenti in loco e li indirizzino verso le risorse di calcolo più appropriate all'interno del sistema a più livelli, secondo le necessità
- I livelli devono fornire risorse di calcolo che includono tutte le risorse, i componenti e i servizi necessari per permettere agli utenti finali di sfruttare la potenza di calcolo; di conseguenza, devono funzionare come pool di esperti nazionali per il supporto agli utenti e ai ricercatori, anche per gli sviluppi dell'IA.

In tal modo il vasto panorama accademico svizzero potrà disporre di un'infrastruttura di calcolo completa, adeguatamente dimensionata, sostenibile e resiliente, in grado di soddisfare un'ampia gamma di esigenze. Il CSS si prefigura la creazione di un sistema di infrastrutture di calcolo a più livelli che assicuri una maggiore efficienza grazie all'interoperabilità, che fornisca agli utenti una transizione fluida tra i livelli, p. es. per l'apprendimento federato, e che sia gestito attraverso un unico punto di accesso (p. es. per i livelli 0, 1 e 2), tenendo conto delle diverse fonti di finanziamento e delle strutture di governance già esistenti.

52 Esempio: CSCS: link all'infrastruttura internazionale di supercalcolo LUMI (Finlandia); supercomputer nazionale: Alps.

53 In questo contesto, il termine «supercalcolo» fa riferimento al calcolo di fascia più alta, ovvero ai sistemi più avanzati che operano ai limiti estremi della tecnologia attuale.

54 Accesso a LUMI: <https://lumi-supercomputer.eu/get-started-2021/users-in-switzerland/>. Per le attività di Helvetic AI Resources, Technologies and Services (HEARTS): <https://csc.fi/en/news/eurohpc-ju-selects-ai-factory-antennas-to-broaden-ai-factories-initiative/> e <https://ggbaweb.ch/en/switzerland-joins-european-ai-factory-network-with-new-hearts-antenna/> (consultati il 26 gennaio 2026).

55 Esempio: infrastrutture di calcolo ad alte prestazioni centrali (soluzioni on-premise) di università cantionali o di scuole universitarie professionali (SUP).

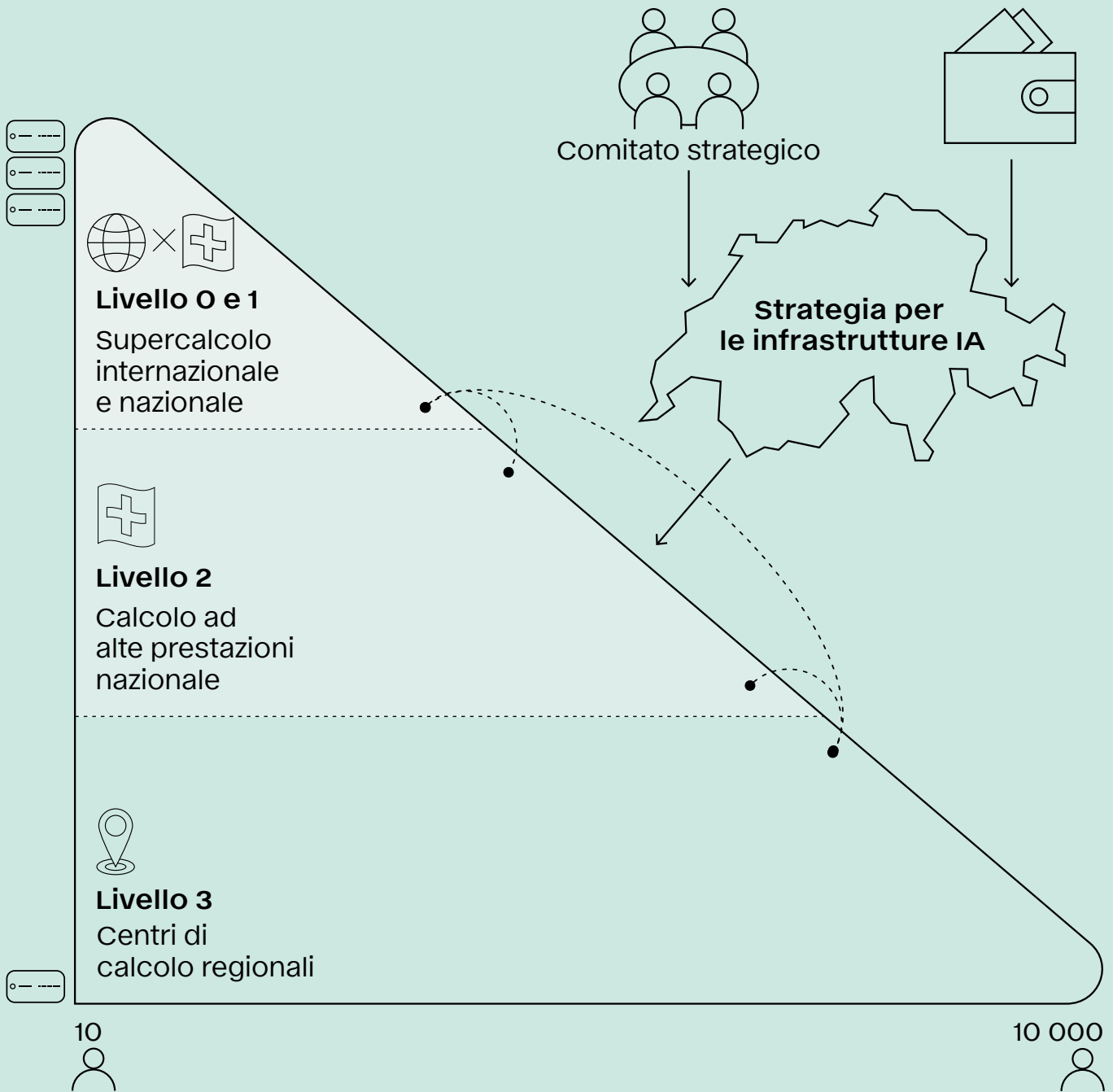


Figura 4: Illustrazione di un sistema di infrastruttura di calcolo a più livelli. L'asse X illustra il numero degli utenti del sistema, mentre l'asse Y mostra il livello di capacità di calcolo. Il numero di utenti mostrato sull'asse X non rappresenta il numero effettivo, ma è solo a scopo illustrativo. Le linee tratteggiate rappresentano la permeabilità dei singoli livelli. I semicerchi tratteggiate rappresentano l'interoperabilità. Qui, l'interoperabilità indica la capacità di scambiare e utilizzare informazioni a tutti i livelli, p. es. dati (compreso il software) e capacità di calcolo.

## 2. Comitato strategico agile per l'infrastruttura di calcolo

- Istituire un comitato strategico nazionale, indipendente e neutrale che valuti la progettazione di un'infrastruttura di calcolo orientata al futuro sotto forma di un sistema a più livelli e prenda decisioni in merito. Tale comitato è responsabile di:
  - progettare strategicamente il sistema a più livelli, sulla base di una continua valutazione delle esigenze nazionali in materia di risorse di calcolo su larga scala e degli sviluppi internazionali delle infrastrutture di calcolo, tenendo conto della governance multilivello. La strategia nazionale deve poggiare sui principi guida sopra descritti.
  - monitorare l'attuazione e l'efficacia della strategia nazionale da parte di ciascun livello del sistema di infrastruttura di calcolo a più livelli.

Il Consiglio federale dovrebbe nominare questo comitato strategico come commissione indipendente di esperti, dotata dei poteri necessari per adempiere alle proprie responsabilità, se necessario su base giuridica.

- Il comitato strategico dovrebbe riunire tutte le competenze pertinenti per adempiere alle proprie responsabilità, comprese quelle tecniche<sup>56</sup>, scientifiche<sup>57</sup>, strategiche<sup>58</sup>, giuridiche<sup>59</sup> e finanziarie<sup>60</sup>, rappresentando al contempo gli utenti e i gruppi di interesse delle risorse di calcolo nazionali

e regionali delle scuole universitarie. La dimensione del comitato deve consentire un orientamento strategico efficiente.

## 3. Implicazioni finanziarie

- Adeguare i finanziamenti in modo da coprire integralmente le risorse di calcolo; il finanziamento può contare solo in parte sulle sovvenzioni, poiché queste coprono solo i costi sul breve periodo. Le risorse di calcolo, tuttavia, richiedono un impegno e investimenti a lungo termine, per cui è necessario assicurare per tutti i livelli un finanziamento adeguato e duraturo. In generale, le risorse di calcolo devono essere rese accessibili in modo trasparente ed economicamente efficiente, il che significa che i costi devono essere coperti ma non si devono generare profitti a carico delle scuole universitarie.

56 Competenze tecniche: architetture hardware avanzate, acceleratori e chip specializzati, software di sistema e programmazione, ottimizzazione e messa a punto delle prestazioni, infrastrutture di rete e dati, gestione dei dati e tecnologie emergenti (p. es. calcolo quantistico).

57 Competenze scientifiche: competenza di ricerca specializzata per settore: comprensione delle esigenze specifiche in materia di dati, calcolo e simulazione di varie aree di ricerca per garantire che la progettazione di sistemi di calcolo ad alte prestazioni sia in linea con le esigenze scientifiche concrete.

58 Competenze strategiche: collaborazione interdisciplinare e lungimiranza strategica: promuovere la cooperazione tra esperti tecnici, scientifici, politici ed economici, mantenendo al contempo una visione strategica a lungo termine per allineare l'infrastruttura di calcolo alle tendenze emergenti della ricerca e del mercato.

59 Competenze giuridiche: dinamiche dell'ecosistema e della catena di approvvigionamento: comprensione delle relazioni tra fornitori, considerazioni in materia di proprietà intellettuale e conformità ai quadri normativi per mitigare i rischi e garantire un'infrastruttura di calcolo giuridicamente solida, inclusi i sistemi di memorizzazione e archiviazione associati. Gestione dei rischi e conformità normativa: gestire i rischi a livello legale e di sicurezza informatica, compresa l'osservanza dei controlli sulle esportazioni, delle normative sulla privacy dei dati e di altri obblighi giuridici, al fine di garantire che gli investimenti strategici siano resilienti e conformi.

60 Competenze finanziarie: analisi strategica degli investimenti e acume finanziario: capacità di condurre analisi costi-benefici approfondite e di gestire strategie d'investimento a lungo termine, bilanciando prestazioni, scalabilità e responsabilità fiscale.

# 1 Framing and context

## 1.1 Why this topic – AI and computing infrastructure?

With the deployment of large language models (LLM) like ChatGPT in late 2022, the discussions around AI have reached a new high as it has been shown how it can support the work of many. The rapid technological development of AI, which goes beyond LLM, touches on many aspects of society, including education, research and innovation. AI will have a significant impact on various aspects of research, including the work of researchers, the research process, the publication process, the fundamental principles that underpin scientific work, and the requirements for computing infrastructures and private-public partnerships.<sup>61,62,63,64,65,66,67,68</sup>

The reason for addressing the topic of ‘AI and computing infrastructure’ is that research involving large amounts of data has become increasingly important over the last decade. AI methods, such as machine learning, offer extensive opportunities in data-driven research and have certainly fuelled this trend. Access to adequate computing infrastructure is essential for AI-driven, data- and compute-intensive analyses. The SSC therefore wanted to establish whether Swiss academia requires a publicly funded AI computing infrastructure or whether computing resource services provided by industry (e.g., hyperscalers<sup>69</sup>) are sufficient. The rationale was that public funding for the Alps supercomputer<sup>70</sup> had been planned for the current ERI period (2025–2028<sup>71</sup>) and previous ERI periods (e.g., 2021–2024<sup>72</sup>, 2017–2020<sup>73</sup>),

- 61 Scientific Advice Mechanism. Successful and timely uptake of artificial intelligence in science in the EU: Evidence review report. Berlin: SAPEA. 2024. doi:10.5281/zenodo.10849580. <https://scientificadvice.eu/advice/artificial-intelligence-in-science/> (last checked on 26.01.26).
- 62 Maslej, N., Fattorini, L., Brynjolfsson, E., Etchemendy, J., Ligett, K., Lyons, T., Manyika, J., Ngo, H., Niebles, J.C., Parli, V., Shoham, Y., Wald, R., Clark, J. and Perrault, R. “The AI Index 2023 Annual Report,” AI Index Steering Committee, Institute for Human-Centered AI, Stanford University, Stanford, CA, April 2023.
- 63 Maslej, N., Fattorini, L., Perrault, R., Parli, V., Reuel, A., Brynjolfsson, E., Etchemendy, J., Ligett, K., Lyons, T., Manyika, J., Niebles, J.C., Shoham, Y., Wald, R. and Clark, J. “The AI Index 2024 Annual Report,” AI Index Steering Committee, Institute for Human-Centered AI, Stanford University, Stanford, CA, April 2024.
- 64 Group of Chief Scientific Advisors. Successful and timely uptake of artificial intelligence in science in the EU: Independent expert report. Scientific Opinion No. 15. 2024. doi:10.2777/08845. <https://scientificadvice.eu/advice/artificial-intelligence-in-science/> (last checked on 26.01.26).
- 65 Maslej, N., Fattorini, L., Perrault, R., Gil, Y., Parli, V., Kariuki, N., Capstick, E., Reuel, A., Brynjolfsson, E., Etchemendy, J., Ligett, K., Lyons, T., Manyika, J., Niebles, J.C., Shoham, Y., Wald, R., Walsh, T., Hamrah, A., Santarlasci, L., Betts Lotufo, J., Rome, A., Shi, A., Oak, S. “The AI Index 2025 Annual Report,” AI Index Steering Committee, Institute for Human-Centered AI, Stanford University, Stanford, CA, April 2025. <https://doi.org/10.48550/arXiv.2504.07139> (last checked on 26.01.26).
- 66 Purificato, E., Bili, D., Jungnickel, R., Ruiz Srna, V., Fabiani, J. et al. The Role of Artificial Intelligence in Scientific Research – A Science for Policy, European Perspective, Publications Office of the European Union, Luxembourg, 2025. <https://data.europa.eu/doi/10.2760/7217497> (last checked on 26.01.26), JRC143482.
- 67 European Commission: Directorate-General for Research and Innovation, Campbell, D., Iversen, E., Karlström, H., Labrosse, I. et al. The use of generative artificial intelligence in research, Publications Office of the European Union, 2025. <https://data.europa.eu/doi/10.2777/1024414> (last checked on 26.01.26).
- 68 Swiss young academy. Impact of AI on Early Career Researchers: Challenges, Opportunities and Responsibilities. December 2025. <https://de.swissyoungacademy.ch/publications/impact-of-ai-on-early-career-researchers-challenges-opportunities-and-responsibilities> (last checked on 26.01.26).
- 69 Refers to cloud service providers (such as AWS, Google, Microsoft or other companies) that offer hyperscale data centre services.
- 70 See <https://www.cscs.ch/computers/alps> (last checked on 26.01.26).
- 71 HPCN-28: see SBFI. 24.031. Botschaft zur Förderung von Bildung, Forschung und Innovation in den Jahren 2025–2028. BBI 2024 900. 8. März 2024. See <https://www.fedlex.admin.ch/eli/fga/2024/900/de> and SBFI. Schweizer Roadmap für Forschungsinfrastrukturen im Hinblick auf die BFI-Botschaft 2025–2028 (Roadmap Forschungsinfrastrukturen 2023). Teil I: Nationale Forschungsinfrastrukturen. SBFI 2023. See <https://www.sbf.admin.ch/de/schweizer-roadmap-fuer-forschungsinfrastrukturen> (last checked on 26.01.26).
- 72 HPCN-24: see SBFI. 20.028. Botschaft zur Förderung von Bildung, Forschung und Innovation in den Jahren 2021–2024. BBI 2020 3681. 5. Mai 2020. See <https://www.fedlex.admin.ch/eli/fga/2020/866/de> and SBFI. Schweizer Roadmap für Forschungsinfrastrukturen im Hinblick auf die BFI-Botschaft 2021–2024 (Roadmap Forschungsinfrastrukturen 2019). [https://www.sbf.admin.ch/dam/de/sd-web/3VmrDpFwm7YL/roadmap\\_forschungsinfrastrukturen\\_2019\\_de.pdf](https://www.sbf.admin.ch/dam/de/sd-web/3VmrDpFwm7YL/roadmap_forschungsinfrastrukturen_2019_de.pdf) (last checked on 26.01.26).
- 73 HPCN-20: see SBFI. 16.025. Botschaft zur Förderung von Bildung, Forschung und Innovation in den Jahren 2017–2020. BBI 2016 3089. 24. Februar 2016. See <https://www.fedlex.admin.ch/eli/fga/2016/576/de> and SBFI Schweizer Roadmap für Forschungsinfrastrukturen im Hinblick auf die BFI-Botschaft 2017–2020 (Roadmap Forschungsinfrastrukturen 2015). [https://www.sbf.admin.ch/dam/de/sd-web/dn93p8fHHZPH/roadmap\\_forschungsinfrastrukturen\\_2015\\_de.pdf](https://www.sbf.admin.ch/dam/de/sd-web/dn93p8fHHZPH/roadmap_forschungsinfrastrukturen_2015_de.pdf) (last checked on 26.01.26).

and the Swiss AI Initiative<sup>74</sup> has been launched. This raises the question of whether these initiatives should continue to be funded during the next ERI period (2029–2032) because they benefit academic research in Switzerland, or whether there are preferable solutions in terms of benefits, risks and costs. Secondly, to the knowledge of the SSC, there has been no cross-academy assessment of the computing resources available for AI research in Swiss academia. It is therefore not clear to what extent the Swiss academic landscape has a computing infrastructure that is sufficient for data- and compute-intensive analyses and to what extent such an infrastructure is needed. Another consideration is that other countries appear to be struggling to retain top AI talent in academia due to a brain drain to industry.<sup>75</sup> It is essential to have sufficient access to technology experts in academia who can help researchers explore and make good use of the computing infrastructure and associated AI tools.<sup>76</sup> Whether the brain drain also applies to Switzerland, is an additional open question.

## 1.2 What currently exists in terms of AI and computing infrastructure

### 1.2.1 International computing infrastructure

The EU has an AI Continent Action Plan<sup>77</sup>, which includes beyond other actions building large-scale AI data and computing infrastructures, increasing access to high-quality data and fostering AI skills. AI Factories with associated antennas<sup>78</sup> and AI Gigafactories<sup>79</sup> are established. The EuroHPC Joint Undertaking (JU)<sup>80</sup> provides access to the computing time and support services offered by the EuroHPC AI Factories<sup>81</sup>, which are complemented by AI antennas. The AI Factories are open to European users from various sectors, including industry, research, academia and public authorities. The LUMI AI Factory<sup>82</sup> is one example of the EuroHPC JU initiative, that pools EU supercomputing resources and services. LUMI<sup>83</sup>, located in Finland, is the associated pre-exa-scale supercomputer, which is hosted by the LUMI consortium<sup>84</sup>. LUMI is co-funded by EuroHPC JU and the LUMI consortium. The LUMI AI Factory has a LUMI AI Factory Consortium<sup>85</sup> that is formed by six countries<sup>86</sup> of the LUMI consortium. The LUMI AI Factory is hosted by CSC<sup>87</sup>, which is a non-profit limited liability company,

74 See <https://www.swiss-ai.org/> (last checked on 02.12.25).

75 Maslej, N., Fattorini, L., Perrault, R., Parli, V., Reuel, A., Brynjolfsson, E., Etchemendy, J., Ligett, K., Lyons, T., Manyika, J., Niebles, J.C., Shoham, Y., Wald, R. and Clark, J. "The AI Index 2024 Annual Report," AI Index Steering Committee, Institute for Human-Centered AI, Stanford University, Stanford, CA, April 2024.

76 The Alan Turing Institute. Technopolis group. Review of Digital Research Infrastructure Requirements for AI. 2022. See <https://www.turing.ac.uk/news/publications/review-digital-research-infrastructure-requirements-ai> (last checked on 26.01.26).

77 See <https://digital-strategy.ec.europa.eu/en/policies/european-approach-artificial-intelligence> (last checked on 26.01.26).

78 AI antennas are associated to AI-optimised supercomputers in existing AI Factories; see <https://digital-strategy.ec.europa.eu/en/policies/ai-factories> (last checked on 26.01.26).

79 AI Gigafactories are large facilities for the development and training of next-generation AI models with trillions of parameters. They are expected to host computing power of 100,000 advanced AI processors. See <https://digital-strategy.ec.europa.eu/en/policies/ai-factories> (last checked on 26.01.26).

80 See <https://digital-strategy.ec.europa.eu/en/policies/high-performance-computing-joint-undertaking> (last checked on 26.01.26).

81 See [https://www.eurohpc-ju.europa.eu/ai-factories\\_en](https://www.eurohpc-ju.europa.eu/ai-factories_en) (last checked on 26.01.26).

82 See <https://lumi-ai-factory.eu/> (last checked on 26.01.26).

83 See <https://lumi-supercomputer.eu/about-lumi/> (last checked on 26.01.26).

84 LUMI consortium countries are Finland, Belgium, Czechia, Denmark, Estonia, Iceland, the Netherlands, Norway, Poland, Sweden and Switzerland. See <https://lumi-supercomputer.eu/about-lumi/> (last checked on 26.01.26).

85 See <https://lumi-ai-factory.eu/about-us/lumi-aif-consortium/> (last checked on 26.01.26).

86 The six countries are: Czechia, Denmark, Estonia, Finland, Norway and Poland; see <https://lumi-ai-factory.eu/about-us/lumi-aif-consortium/> (last checked on 26.01.26).

87 See <https://csc.fi/en/about-us/> (last checked on 26.01.26).

owned by the Finnish Government and Finnish higher education institutions. CSC's mission is to build digital solutions for research, national education, culture and public administration in Finland, while also supporting companies. JUPITER<sup>88</sup> is an example of an exascale supercomputer that is part of EuroHPC JU, which is located in Germany.<sup>89</sup> JUPITER was built by a supercomputer consortium consisting of ParTec and Eviden and was procured by EuroHPC JU in collaboration with the Jülich Supercomputing Centre (JSC).<sup>90</sup> Half of JUPITER was funded by EuroHPC JU and the other half by the German Federal Ministry of Research, Technology and Space and the Ministry of Culture and Science of the German state of North Rhine-Westphalia through the Gauss Centre for Supercomputing (GCS).<sup>91</sup> In Germany, the GCS, a non-profit organisation, combines the three national supercomputing centres, the High Performance Computing Center Stuttgart (HLRS), JSC, and Leibniz Supercomputing Centre, Garching (LRZ).<sup>92</sup> By being the German member of the Partnership for Advanced Computing in Europe (PRACE), GCS is also involved internationally. Nationally the GCS is part of the Gauss Allianz<sup>93</sup>, the German HPC partnership that coordinates HPC-related activities at the state level.<sup>94</sup>

The EU AI Continent action plan is complemented by the EU Apply AI Strategy and the EU AI in Science Strategy.<sup>95</sup> The Apply AI Strategy aims to drive widespread adoption of trustworthy AI across key industries and the public sector. The AI in Science Strategy aims to position the EU as a hub for AI-driven scientific innovation. The Resource for AI Science in Europe (RAISE)<sup>96</sup>, one key element of the AI in Science Strategy, will be a virtual institute with the goal to pool and coordinate AI resources for developing AI and applying it in science. AI resources will include computing power, data, talent and research funding across the EU and private sector. The EU Data Union Strategy<sup>97</sup> aims to scale up access to high-quality data for AI development, streamline data rules and safeguard the EU's data sovereignty. RI-SCALE<sup>98</sup> is an example of a current Horizon Europe funded project that aims to co-design, prototype and validate a Data Exploitation Platform (DEP) technology that supports scalable computational platforms for AI-driven data exploitation.

88 See <https://www.fz-juelich.de/de/jupiter> (last checked on 26.01.26).

89 See [https://www.eurohpc.europa.eu/supercomputers/our-supercomputers\\_en](https://www.eurohpc.europa.eu/supercomputers/our-supercomputers_en) (last checked on 26.01.26).

90 See <https://www.fz-juelich.de/de/jupiter> (last checked on 26.01.26).

91 See <https://www.gauss-centre.eu/news/jsc-officially-welcomes-jupiter-europes-most-powerful-supercomputer> (last checked on 26.01.26).

92 See <https://www.gauss-centre.eu/about-us> (last checked on 26.01.26).

93 See <https://gauss-allianz.de/en/> (last checked on 26.01.26).

94 See <https://gauss-allianz.de/en/#kompetenznetzwerk> (last checked on 26.01.26).

95 See [https://ec.europa.eu/commission/presscorner/detail/en/ip\\_25\\_2299](https://ec.europa.eu/commission/presscorner/detail/en/ip_25_2299) (last checked on 26.01.26).

96 See [https://research-and-innovation.ec.europa.eu/strategy/strategy-research-and-innovation/our-digital-future/european-ai-science-strategy\\_en](https://research-and-innovation.ec.europa.eu/strategy/strategy-research-and-innovation/our-digital-future/european-ai-science-strategy_en) (last checked on 26.01.26).

97 See <https://digital-strategy.ec.europa.eu/en/policies/data-union> (last checked on 26.01.26).

98 See <https://www.riscale.eu/> and <https://www.egi.eu/project/ri-scale/> (last checked on 26.01.26).

Private actors such as Nvidia are also advancing AI factory deployments in Europe partnering with European institutions and companies.<sup>99,100</sup> Worldwide, Nvidia is deploying over 80 new AI-powered scientific systems.<sup>101</sup> In addition, Nvidia is driving a fundamental redesign of the hardware infrastructure for AI. These designs are moving towards rack-scale systems, high voltage direct current distribution and advanced liquid cooling to support hundreds of kilowatts per rack or even megawatt racks in the near future for next-generation AI facilities.<sup>102</sup>

The AI supercomputer Gefion in Denmark is another example of a public-private partnership. It is owned and operated by the Danish Centre for AI Innovation (DCAI)<sup>103</sup>, a company funded by the Novo Nordisk Foundation and the Export and Investment Fund of Denmark (EIFO)<sup>104, 105, 106, 107</sup>. The DCAI works with customers from academia, start-ups and enterprises with the aim to accelerate AI research and innovation.

As an example of a non-EU member state, the UK has an AI and Opportunities Action Plan<sup>108</sup>, to which the UK government responded.<sup>109</sup> The plan and the response outline strategic milestones, such as building sufficient, secure and sustainable AI infrastructure with a compute strategy, creating AI growth zones, forming an AI energy council and creating an AI skills and talent pipeline. The milestones are associated with a substantial budget. The

AI Research Resource (AIRR) provides AI-specialised supercomputing capacity to support researchers, academia and industry.<sup>110</sup> The current AIRR supercomputers are Isambard-AI<sup>111</sup> at Bristol University and Dawn<sup>112</sup> at the University of Cambridge.

In the USA, the Genesis Mission<sup>113</sup> is a U.S. national initiative to harness AI for accelerating scientific breakthroughs by integrating supercomputers, AI systems, experimental facilities (e.g., quantum technologies), data assets and public-private collaboration into a coordinated AI experimentation platform.<sup>114</sup> The U.S. Department of Energy has announced new AMD-accelerated AI supercomputers built under a public-private partnership to accelerate scientific discovery and maintain U.S. leadership in AI computing.<sup>115</sup>

The International Computation and AI Network (ICAIN) aims to rebalance global access to AI computing power. It connects supercomputing resources worldwide to support the UN Sustainable Development Goals, enabling equitable AI research in areas such as climate, health and sustainable growth.<sup>116</sup>

The international computing infrastructures in the field of AI presented here are examples and therefore not exhaustive.

99 See <https://blogs.nvidia.com/blog/gtc-paris-2025/> and <https://nvidianews.nvidia.com/news/nvidia-builds-worlds-first-industrial-ai-cloud-to-advance-european-manufacturing> (last checked on 26.01.26).

100 See <https://www.intelligentcio.com/eu/2025/11/25/ai-factory-france-supercharges-ai-adoption-by-french-start-ups-and-smes-with-nvidia/> (last checked on 26.01.26).

101 See <https://blogs.nvidia.com/blog/sc25-new-science-systems-worldwide/> (last checked on 26.01.26).

102 See <https://developer.nvidia.com/blog/nvidia-800-v-hvdc-architecture-will-power-the-next-generation-of-factories/> (last checked on 26.01.26).

103 See <https://dcai.dk/> (last checked on 26.01.26).

104 See <https://www.eifo.dk/en/> (last checked on 26.01.26).

105 See <https://novonordiskfonden.dk/en/news/denmarks-first-ai-supercomputer-is-now-operational/> (last checked on 26.01.26).

106 See <https://dcai.dk/news/denmark-s-new-ai-supercomputer-geffion-ranked-as-7th-fastest-storage-systems-in-the-world> (last checked on 26.01.26).

107 See <https://blogs.nvidia.com/blog/denmark-sovereign-ai-supercomputer/> (last checked on 26.01.26).

108 See <https://www.gov.uk/government/publications/ai-opportunities-action-plan/ai-opportunities-action-plan> (last checked on 26.01.26).

109 See [https://assets.publishing.service.gov.uk/media/678639913a9388161c5d2376/ai\\_opportunities\\_action\\_plan\\_government\\_reponse.pdf](https://assets.publishing.service.gov.uk/media/678639913a9388161c5d2376/ai_opportunities_action_plan_government_reponse.pdf) (last checked on 26.01.26).

110 See <https://www.gov.uk/government/publications/ai-research-resource#full-publication-update-history> (last checked on 26.01.26).

111 See <https://www.bristol.ac.uk/research/centres/bristol-supercomputing/articles/2025/isambard-ai-optimised-supercomputer.html> (last checked on 26.01.26).

112 See <https://www.cam.ac.uk/stories/under-the-bonnet-at-AI-supercomputer-Dawn> (last checked on 26.01.26).

113 See <https://genesis.energy.gov/> (last checked on 26.01.26).

114 Gibney, E., Witze, A., Ahart, J.. Trump's AI 'Genesis Mission': what are the risks and opportunities? Nature. 2025;648(8093):253-255. doi:10.1038/d41586-025-03890-z.

115 See <https://www.energy.gov/articles/energy-department-announces-new-public-private-partnership-model-two-supercomputers> (last checked on 26.01.26).

116 See <https://icain.ch/> (last checked on 26.01.26).

## 1.2.2 National computing infrastructure

The Swiss National Supercomputing Centre (CSCS)<sup>117</sup>, which is part of ETHZ, hosts the Alps supercomputer, which became operational in 2024 and provides large-scale GPU resources for AI model training and scientific applications. On behalf of the Swiss Confederation, CSCS operates the User lab in the scope of the national initiative for High Performance Computing and Networking (HPCN), thereby providing researchers access to this national supercomputer and to international supercomputing. The supercomputing resources are financed by the HPCN initiative, as outlined in the Swiss Roadmap for Research Infrastructures, and through dedicated funding of the ERI Dispatch<sup>118</sup>. The operational costs are covered by ETHZ.<sup>119</sup> Researchers can apply for computing time using different allocation schemes.<sup>120</sup> Allocation is provided upon positive review.<sup>121</sup> Because Switzerland is a member of the LUMI consortium (see 1.2.1), Swiss researchers can also apply for computing time on LUMI via calls of CSCS.<sup>122</sup> The Helvetic AI Resources, Technologies and Services (HEARTS), which includes the Alps research infrastructure, was selected as AI Factory antenna of the EuroHPC JU. HEARTS will work closely with the AI Factory LUMI AI in Finland and will collaborate with the BSC AI Factory in Spain, the MIMER AI Factory in Sweden and the IT4LIA AI Factory in Italy. It will focus pri-

marily on meteorology, environmental modelling and AI-optimised data services and will serve national and European users across academia, start-ups, SME and the public sector.<sup>123</sup>

The current set-up or governance of the CSCS might change with the ETH domain restructuring project, 'Fit for the Future'. It envisions the creation of a Digital Unit as a mission-oriented unit that would transfer the CSCS and the Swiss Data Science Center (SDSC) under the same roof that was proposed for the ETH domain research institutes.<sup>124</sup>

In December 2023, ETHZ and EPFL launched the Swiss AI Initiative, supported by CSCS, with an allocation of 10 million GPU hours on Alps and 20 million CHF in funding from the ETH Domain.<sup>125</sup> The initiative is the first initiative of the Swiss National AI Institute (SNAI), a partnership between the ETH AI Center and the EPFL AI Center, and focuses on the development of open-source foundation models.<sup>126</sup> Apertus is a multilingual large language model developed as part of the Swiss AI Initiative.<sup>127,128</sup>

Switch is a Swiss foundation that provides secure digital infrastructure and services for national universities and research institutions and links them to international networks, such as the Gigabit European Academic Network (GÉANT)<sup>129</sup> or the European Open Science Cloud (EOSC) Association.<sup>130,131</sup> Switch supports academic research by offering networking, cloud, identity manage-

117 See <https://www.cscs.ch/> (last checked on 26.01.26).

118 SERI. 24.031. Dispatch on the promotion of Education, Research and Innovation in the years 2025–2028. BBI 2024 900. 8 March 2024. See <https://www.fedlex.admin.ch/eli/fga/2024/900/de> (last checked on 26.01.26; text not available in English).

119 SERI. Swiss Roadmap for Research Infrastructures with a view to the ERI Dispatch 2025–2028 (Roadmap Research Infrastructures 2023). Part I: National Research Infrastructures. SERI 2023. See <https://www.sbf.admin.ch/en/swiss-roadmap-for-research-infrastructures> (last checked on 26.01.26).

120 See <https://www.cscs.ch/user-lab/allocation-schemes> (last checked on 26.01.26).

121 See <https://www.cscs.ch/user-lab/overview> (last checked on 26.01.26).

122 See <https://lumi-supercomputer.eu/get-started-2021/users-in-switzerland/> (last checked on 26.01.26).

123 See <https://csc.fi/en/news/eurohpc-ju-selects-ai-factory-antennas-to-broaden-ai-factories-initiative/> and <https://ggbaswiss.ch/en/switzerland-joins-european-ai-factory-network-with-new-hearts-antenna/> (last checked on 26.01.26).

124 See <https://ethrat.ch/en/eth-domain/objectives-of-fit-for-the-future-and-decision-in-principle/> (last checked on 26.01.26).

125 See <https://www.swiss-ai.org/> (last checked on 26.01.26).

126 See <https://ethz.ch/en/news-and-events/eth-news/news/2024/10/eth-zurich-and-epfl-enhance-collaboration-to-boost-ai-in-switzerland.html> (last checked on 26.01.26).

127 See <https://www.swiss-ai.org/apertus> (last checked on 26.01.26).

128 See [https://ethz.ch/en/news-and-events/eth-news/news/2025/09/press-release-apertus-a-fully-open-transparent-multilingual-language-model.html#:~:text=Apertus%20was%20developed%20as%20part,National%20Supercomputing%20Centre%20\(CSCS\)](https://ethz.ch/en/news-and-events/eth-news/news/2025/09/press-release-apertus-a-fully-open-transparent-multilingual-language-model.html#:~:text=Apertus%20was%20developed%20as%20part,National%20Supercomputing%20Centre%20(CSCS)) (last checked on 26.01.26).

129 GÉANT is the collaboration of European National Research and Education Networks (NRENs). See <https://geant.org/> (last checked on 26.01.26).

130 The EOSC Association coordinates Europe's federated research data infrastructure. Switch is member of EOSC Association. See <https://eosc.eu/members/switch/> (last checked on 26.01.26).

131 Switch is one of 38 pan-european project partners in the collaboration of NRENs. See <https://geant.org/projects/partners/> (last checked on 26.01.26).

ment and computing resources with the aim to simplify access to advanced digital tools.<sup>132,133</sup> Switch is financed as a non-profit foundation, established by the Swiss Confederation and university cantons, and sustained through service revenues and stakeholder contributions without profit maximation.<sup>134,135</sup>

In addition to these public computing capacity providers that serve the Swiss ERI landscape, there are also private providers. Examples include hyperscalers such as AWS, Microsoft and Google, as well as Exoscale, among others. Exoscale is a Swiss-based European cloud provider that supports academic research by offering secure, scalable infrastructure services, while also contributing to international scientific collaboration through its integration into European research networks, e.g., GÉANT.<sup>136,137</sup> It is financed through commercial operations and is a member of AI Digital, which is part of the AI Telekom Austria Group.<sup>138,139</sup>

Several Swiss organisations are participating in ICAIN.<sup>140</sup> The founding members include the Swiss Federal Department of Foreign Affairs (FDFA), ETHZ, CSCS, EPFL as well as international organisations such as the European Laboratory for Learning and Intelligent Systems (ellis)<sup>141</sup>, Data Science Africa (DSA)<sup>142</sup> and CSC.

The national computing infrastructures presented here are examples and therefore not exhaustive.

132 See <https://www.switch.ch/en/insights/switch-cloud-digital-sovereignty-universities> (last checked on 26.01.26).

133 Switch. Strategy 2025+. See [https://static.www.switch.ch/sites/default/files/2025-09/switch-strategy-2025v1.0\\_en\\_0.pdf](https://static.www.switch.ch/sites/default/files/2025-09/switch-strategy-2025v1.0_en_0.pdf) (last checked on 26.01.26).

134 See <https://www.switch.ch/en/about/our-foundation> (last checked on 26.01.26).

135 Switch Annual Report 2024. See [https://static.www.switch.ch/sites/default/files/2025-06/Switch\\_Annual\\_Report\\_2024\\_EN.pdf](https://static.www.switch.ch/sites/default/files/2025-06/Switch_Annual_Report_2024_EN.pdf) (last checked on 26.01.26).

136 See <https://www.vshn.ch/ocre/> (last checked on 26.01.26).

137 See <https://www.exoscale.com/ocre/> (last checked on 26.01.26).

138 See <https://www.ai.digital/european-cloud/> (last checked on 26.01.26).

139 See <https://www.exoscale.com/lp/swiss-cloud-hosting/> and <https://www.exoscale.com/about-us/> (last checked on 26.01.26).

140 See <https://icain.ch/#founding-partners-part> (last checked on 26.01.26).

141 See <https://ellis.eu/> (last checked on 26.01.26).

142 See <https://www.datascienceafrica.org/> (last checked on 26.01.26).

## 1.3 Approach to address AI and computing infrastructure

The Council conducted exploratory interviews with thirteen national organisations in the higher education sector (including cantonal universities, universities of applied sciences (UAS), the ETH domain and other organisations linked to the ERI domain), as well as four international supercomputer centres. The aim was threefold. First, it intended to determine the computing capacities currently available or utilised within the HEI for data- and compute-intensive research (i.e. the current needs). Secondly, to identify the potential future changes to these capacities within the next five to ten years, such as expanding or reducing in-house computing resources or industry-provided services. The focus was not only on the current situation, but also on the organisations' visions for the future design of such computing infrastructures, with the aim of identifying best practices and ideas on this topic. Thirdly, to ascertain whether Switzerland, like other countries, is experiencing difficulties in retaining highly qualified AI experts in academic research, and if so, the reasons for this. The national HEI were interviewed using the national interview guide for determining computing infrastructure resources (see Appendix 4.1.1). To test this approach, a pilot study was conducted, in which three HEI from different categories were interviewed using the national interview guide. The pilot study validated the national interview

guide and assessed the feasibility of the approach before the rest of the national organisations were interviewed. International supercomputer centres were interviewed using the international interview guide (see Appendix 4.1.2). All organisations that were interviewed were given the opportunity to complete the interview guides in writing after the initial online interview. The results of the national and international interviews were discussed during several meetings of the AI and Infrastructure working group as well as during several plenary meetings of the entire Council. The recommendations developed were approved by the entire Council during a plenary meeting. To receive feedback on the approved recommendations, five national organisations in the higher education sector were interviewed a second time. They were asked the standardised questions depicted in Annex 4.1.3. The working group approved the adapted recommendations and the accompanying report, which was also approved by the entire Council during a plenary meeting (see Appendix 4.2 for details of the entire working process).

## 2 Reasoning behind the recommendations

### 2.1 National strategy – guiding principles

Switzerland needs to develop a long-term AI infrastructure strategy that serves all public ERI needs. This national strategy has to include a modular computing infrastructure leveraging existing synergies, by forecasting and aligning the needs of the ERI domain's national large-scale computing resources and by enabling Switzerland to participate in the development of international computing infrastructures. It is to outline how a tiered computing infrastructure system meets user requirements for flexibility, scalability and efficiency, including cost-efficiency, while also being resilient and sustainable considering natural resources. It should value digital sovereignty and knowledge security, and foster partnerships with value-sharing stakeholders. It should include a data lifecycle management strategy for the tiered computing infrastructure system that builds on the FAIR<sup>143</sup> principles and it also ought to set out how to promote and sustain world-class AI expertise in Switzerland that can compete internationally.

#### Why?

The guiding principles form the quintessence of the important aspects mentioned in the interviews regarding a computing infrastructure that enables the use of AI for ERI in Switzerland as well as other academic research requiring intensive computing resources. Several challenges were stated by the interviewed organisations. Predicting resources is challenging as it is difficult to forecast the future need for, and quantity of high-end computing resources, and how these will be affected or reduced by optimisation, e.g., at software level. The democratisation of computing resources is perceived as important as HEI face institutional limitations. This is because single HEI computing resources are not easy to finance sustainably locally, are often insufficient and cannot compare to those provided by industry. Current costs for computing resources make it impossible to have competitive local resources, and computing resources have a short turnover rate, meaning they become outdated quickly. But all Swiss HEI have the need for com-

puting resources as well as data access, storage and archiving. This is why optimisation of computation and funding are suggested as well as implementing a data management strategy. Moving computation closer to data to reduce expensive data transfer and exploring alternative funding mechanisms, beyond financing through grants, to support sustainable computing infrastructure were suggested measures. The SSC recommends counteracting these challenges by developing a long-term AI infrastructure strategy that serves all public ERI needs and that includes a modular computing infrastructure leveraging existing synergies. The Council emphasises that both the recommended strategy and the modular computing infrastructure must be able to adapt to rapid developments in the field of AI, i.e., agility is essential.

Access to existing HEI computing resources is often perceived as complicated and not sufficient for AI research. Thus, the accessibility and adequacy of existing academic computing resources is perceived as suboptimal. Which gives reason to some AI experts to work partly for industry as there sufficient computing resources are provided without question. Some HEI face a high demand for computing resources for teaching, which cannot be financed by grants. Additionally, some HEI experience that principal investigators prefer personal hardware over centralised IT services, which can lead to inefficiencies. Addressing this issue requires a shift in culture, as computing resources encompass not only hardware but also expert personnel, backup systems and disaster recovery plans, elements that cannot be properly maintained at an individual level. At the same time academic competitiveness in research with and on AI is currently being called into question. With industry increasingly leading in AI research, academia risks falling behind and needs to consider how to remain competitive. Having access to sufficient and cost-effective computing resources is suggested as bringing an advantage. Beyond that the Swiss AI Initiative is perceived as primarily benefiting the ETH domain rather than serving as a national initiative with associated highest-end computing resources. AI, however, is relevant to more than just the technical disciplines. International interviewees find that the ability to scale up based on demand is key and envision conversion of cloud computing and high perfor-

143 FAIR: Findable, accessible, interoperable, reusable. Wilkinson, M., Dumontier, M., Aalbersberg, I. et al. The FAIR Guiding Principles for scientific data management and stewardship. *Sci Data* 3, 160018 (2016). FAIR4RS. Barker, M., Chue Hong, N.P., Katz, D.S. et al. Introducing the FAIR Principles for research software. *Sci Data* 9, 622 (2022).

mance computing (HPC), moving away from single super-computers. Thus, a future high-end computing infrastructure must be more flexible and modular as AI distribution will increase. This is why the SSC recommends to create a modular computing infrastructure in form of a tiered system which meets user requirements for flexibility, scalability and efficiency, including cost-efficiency, while also being resilient and sustainable considering natural resources.

Most interviewed organisations mainly use in-house computing resources and provide a central high-performance computing infrastructure, while some also use public cloud and private cloud solutions. Reasons for using in-house computing include concerns about data privacy, protection and sovereignty, as well as predictable costs, as many of these organisations consider hyperscaler cloud computing to be unaffordable. Thus, they are concerned about dependency of and loss of autonomy with external services (e.g., from hyperscalers). Reasons for using cloud computing include complementing in-house computing resources, for example, to absorb peaks in demand or for large-scale use cases, as well as receiving computing resource donations from hyperscalers. The interviewed international organisations provide national, sovereign high-end computing resources to support academic research as well as to industry in varying degrees. Digital sovereignty and data protection are stated as being important also politically. This is why the SSC recommends that the national strategy values digital sovereignty<sup>144</sup> and knowledge security<sup>145</sup> and fosters partnerships with value-sharing stakeholders.

National interviewees find that the questions of data access, storage and archiving are as important as the question of future computing resources and must be discussed at the same time as these questions are correlated. They state that the current challenge is that many initiatives for data storage and/or archiving exist, but that it is not clear what is provided and what is the best solution. This is perceived as a dispersed landscape lacking a clear vision and coordination. Thus, the need for a centralised or federated data storage and archiving sys-

tem that ensures academic research data accessibility nationally was formulated. This is why the SSC recommends that the national strategy should include a data lifecycle management strategy that builds on the FAIR<sup>146</sup> principles, as for the SSC interoperability is key to connect existing repositories.

Retaining AI talent at a senior level is increasingly challenging in the highly competitive job market, which is why most interviewed national higher education institutions experience a brain drain to industry. Difficulties are the lack of job stability in academia as well as the non-competitive salaries. Providing an attractive academic ecosystem, developing an encouraging proper career path that blends academic skills with operational skills needed in AI projects, empowering staff with targeted training programmes, offering meaningful and interesting working content and conditions are mentioned as potential counteracts for brain drain. National interviewees see the need to develop competencies to have top-notch AI expertise in order to negotiate with suppliers and be indispensable for international collaborations and partnerships. International interviewees experience that hosting cutting-edge facilities attracts AI and HPC experts and helps to retain talent. They still experience brain drain to industry, which they counteract by, for example, associating with or being close to excellent universities. For these reasons, the SSC recommends that the national strategy also set out how to promote and sustain world-class AI expertise in Switzerland that can compete internationally.

Even though the SSC refers here to computing infrastructure for the ERI sector, it emphasises that this national strategy should also be aligned with economic strategies so that compatibility and scalability for economic benefit are considered from the outset, for example by allowing SME to carry out their product development on the computing infrastructure recommended here, but with deployment taking place on commercial infrastructures, for example.

144 Swiss Confederation. Digital sovereignty of Switzerland. Report by the Federal Council in response to postulate 22.4411 Z'graggen. 14 December 2022. 26 November 2025. See <https://www.news.admin.ch/en/newnsb/2VPWG78YrV-s4eAVeiklQx> (last checked on 26.01.26; report not available in English).

145 Swissuniversities. Knowledge Security in Switzerland: A Strategic Framework for Higher Education Institutions and Authorities. 15. September 2025. See [https://www.swissuniversities.ch/fileadmin/swissuniversities/Dokumente/Komm/Empfehlungen/Bericht\\_KnowledgeSecurity\\_251127.pdf](https://www.swissuniversities.ch/fileadmin/swissuniversities/Dokumente/Komm/Empfehlungen/Bericht_KnowledgeSecurity_251127.pdf) (last checked on 26.01.26).

146 FAIR: Findable, accessible, interoperable, reusable. Wilkinson, M., Dumontier, M., Aalbersberg, I. et al. The FAIR Guiding Principles for scientific data management and stewardship. *Sci Data* 3, 160018 (2016). FAIR4RS. Barker, M., Chue Hong, N.P., Katz, D.S. et al. Introducing the FAIR Principles for research software. *Sci Data* 9, 622 (2022).

## 2.2 Tiered computing infrastructure system

- To build a tiered computing infrastructure system based on a national strategy and implemented by its stakeholders, guided by the principles outlined above.

### Why?

To address the identified challenges and provide the ERI landscape with an appropriate computing infrastructure over the next five to ten years, the national stakeholders interviewed proposed various solutions: A federal computing infrastructure as a unified national interface, integrating cantonal and federal computing resources and serving as a single point of entry with simple access and fixed allocation of computing resources by default. Repurposing older, highest-end computing infrastructure for non-highest-end users nationally, as computing resources should be allocated according to needs. This means using the most advanced computing resources, i.e. supercomputing, for the most demanding (i.e. computationally intensive) research and higher-end or standard computing resources for less computationally intensive research and student training. A federated computing approach that allows seamless distribution across multiple platforms as traditional high-performance computing and cloud infrastructures converge, particularly due to growing machine learning activity requiring elastic computing capacity. Hybrid computing infrastructures combining local computing resources with public cloud services enhancing flexibility, especially for peak demand. However, it was also stated that cloud computing will complement, but not replace in-house computing resources due to cost and data protection concerns, and that public cloud providers (e.g., CSCS, Switch) are preferred to private options (e.g., hyperscalers). Investing in international computing infrastructure was recognised as necessary, for example through partnerships, as Switzerland cannot further upscale current supercomputing resources due to local natural resource constraints and costs. Some international interviewees mentioned three essential key measures for academic computing infrastructures: 1) having one national solution for all higher education institutions, 2) directly funding the national solution and 3) ensuring the national solution is independent and neutral.

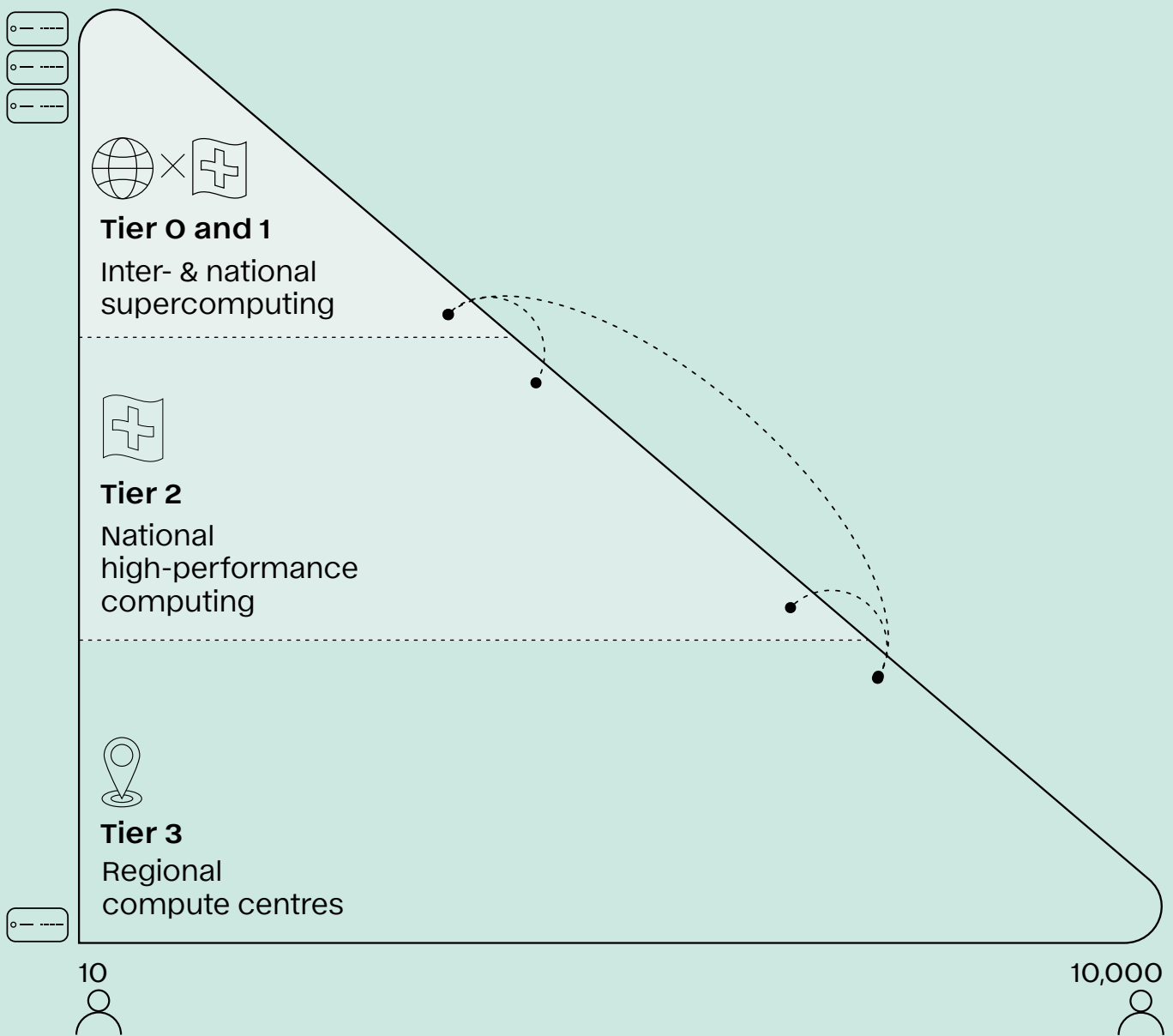


Figure 5: Illustration of the tiered computing infrastructure system. The X-axis exemplifies the number of users of the system, while the Y-axis shows the level of computing capacity. The number of users shown on the X-axis does not represent actual user numbers. It is for illustrative purposes only. The dotted lines represent permeability between the individual tiers. The dotted semicircles represent interoperability. Interoperability here refers to the ability to exchange and use information across all levels, e.g., data (incl. software) and computing capacity.

An appropriate computing infrastructure for the Swiss ERI landscape does not need to be built from scratch, as some components and offerings already exist. CSCS provides national and international researchers with access to the international supercomputer LUMI and the national supercomputer Alps through different allocation schemes. Also, the Swiss AI Initiative runs on Alps. At a local level, HPC infrastructures do exist. Most of the HEI interviewed (e.g., cantonal universities and UAS) provide central HPCs and most of these provide controlled access to their computing resources. Access is generally granted under specific conditions, such as collaboration with internal researchers, affiliation with partner institutions, or via modular payment arrangements. Some initiatives extend sharing to all higher education institutions in the canton or offer services nationwide on a pay-per-use basis. A small subset of interviewed organisations does not share their computing infrastructure. However, these local computing resources are not connected, and some HEI do not provide central HPCs, or the provided computing resources are limited. Furthermore, almost all organisations interviewed will invest in their computing infrastructure within the next five to ten years. This investment is driven by research demand and educational needs. There is a growing demand for machine learning/AI workloads, which require GPU resources. This is a trend that is expected to continue. Scientific computing and data analysis have become essential to almost all fields of science, which is why investment is also being made in classical computing (CPU resources). Additional reasons for investment include lifecycle upgrades and infrastructure renewal. What is currently missing is a national HPC infrastructure that serves a broad range of user needs which do not need highest-end computing resources, and that can be accessed and utilised by all academic researchers in Switzerland under the same conditions.

The SSC recommends establishing a modular computing infrastructure in form of a tiered system. This system comprises three tiers, which differ in terms of computing capacity, number of users, and international, national or regional orientation (see Figure 5). The SSC envisions a tiered computing infrastructure system that enables greater efficiency and scalability by building on interoperability and providing users with seamless transition between the tiers, e.g., for federated learning, and is managed through a single point of access (e.g., for tiers 0, 1 and 2), while taking into account the different funding sources and governance structures already in place. Interoperability here refers to the ability to exchange and use information across all levels, e.g., data (incl. software) and computing capacity. Three considerations were decisive in recommending this tiered computing infrastructure system: 1) democratisation of computing resources is considered important, as all HEI face institutional limitations, and many of the identified challenges cannot be solved by a bottom-up approach alone; 2) almost all organisations will invest in their computing infrastructure in the coming years due to the needs of scientific disciplines; 3) components that can be built upon and expanded in accordance with the guiding principles of the national strategy already exist, so that the recommended tier system to be built will utilise synergies. The SSC is convinced that the outlined system, as detailed below, will provide Switzerland's diverse academic landscape with a comprehensive, appropriately scaled, sustainable and resilient computing infrastructure serving a wide range of needs.

For Tier 0 and 1<sup>147</sup>: International and national supercomputing<sup>148</sup>

- To further develop the existing tiers of national and international supercomputing by expanding the access to and participation in the highest-end international computing resources and integrating them with the national resources.

## Why?

Switzerland is currently well positioned in the field of supercomputing. It has its own national supercomputer, Alps, and access to the international LUMI supercomputer in Finland. However, the turnover rate of these computing capacities is short, raising the question of how Switzerland intends to and can position itself in this regard in the next years. National interviewees stated that no future increase in national supercomputing, referring to highest-end-computing, the most advanced systems operating at the furthest limits of current technology, is possible due to the high energy costs in Switzerland and due to the fact that the necessary energy cannot be delivered. Current trends in redesigning hardware for next-generation AI facilities that require hundreds of kilowatts to megawatts of power and advanced liquid cooling systems might exacerbate this. Thus, local resource constraints, such as the adequate availability of energy for operating such systems and water for cooling them as well as other conditions, such as topological position and access to high-performance connections (i.e. cables) for data transfer, must be taken into account in the design of future supercomputers. The interviewed international supercomputer centres also view space, energy and sustainability considerations, such as available location, potential power capacity and possibilities for heat reuse and cooling, as prerequisites for future scaling of computing resources. However, they either have no problem with these requirements or are willing to pay high energy prices, as digital sovereignty is politically important. Some also find that the ability to scale up computing resources based on demand is key and envision the conversion of cloud computing and HPC, moving away from single supercomputers. Therefore, they conclude that, as AI distribution will increase, a future highest-end computing infrastruc-

ture must be more flexible and modular. National interviewees see two alternatives for a future supercomputing design that take into account the described local constraints: 1) collaborating with hyperscalers at higher costs, or 2) investing in international highest-end computing infrastructures, either by participating in and/or expanding the access to existing highest-end infrastructures such as for example LUMI and EuroHPC, or by building international private-public partnerships in countries with similar innovation potential, sufficient natural resources and favourable topological conditions, while taking digital sovereignty into account. As CERN is located in Switzerland, the countries participating in CERN have high-performance connections to it that could be used for this purpose, i.e. data transfer to a newly built international highest-end computing infrastructure. One proposal was to establish a digitally sovereign giga-computing infrastructure (similar to the EU AI Gigafactories), financed by public-private partnerships, in collaboration with other countries that meet the above requirements. The aim would be for such a giga infrastructure to enable the pre-training of LLM base models. The refinement of the LLM base models could then be carried out by the partner countries in their own countries. This would result in a future basic infrastructure consisting of hardware, software and expertise, i.e., the LLM base models would be part of the software infrastructure and would be pre-trained internationally and refined nationally. The international respondents are also part of an international highest-end computing infrastructure (e.g., EuroHPC) or want to participate and will promote and look for international collaborations. Some of them also mentioned that they are interested in collaborating with or partnering with CSCS and the current national supercomputer Alps. International interviewees also expressed different views on the EU AI Gigafactories. While some are rather critical of them because they are economically oriented and industry driven, with limited academic integration, others see them as the concept of the future.

Taking these considerations into account, the SSC concluded that, due to local constraints, it is essential to further develop the existing tiers of national and international supercomputing. This should be achieved by expanding access to and participation in the highest-end international computing resources, and by integrating these with national resources.

<sup>147</sup> Example: CSCS: link to international supercomputing facility – LUMI (FIN); national supercomputing: Alps.

<sup>148</sup> In this context, 'supercomputing' refers to 'highest-end computing', the most advanced systems operating at the furthest limits of current technology.

- To place the existing national supercomputing provider (CSCS), which also provides access to international supercomputing facilities<sup>149</sup>, under a governance that fulfils its national mandate, to ensure serving all ERI stakeholders throughout Switzerland.

## Why?

CSCS provides researchers with access to national and international supercomputing via the user lab, which is part of the national HPCN Initiative, on behalf of the Swiss Confederation. The user lab is funded by HPCN and operated by ETHZ. Different allocation schemes exist for which researchers can apply. Computing resources received via the User lab upon positive evaluation of the application are free of charge. Some of the interviewed national organisations already profit from the user lab. Others state that it is very difficult to benefit from these resources. Furthermore, although the Swiss AI Initiative with its associated high-end computing resources is perceived as a national initiative, it is seen as primarily benefiting the ETH domain. As AI is relevant for more than the technical disciplines, access to these highest-end computing capacities must be available to all scientific disciplines that require them. Some international interviewees recommended that CSCS should be an independent organisation in Switzerland rather than being under the umbrella of ETHZ. This is because, for a computing infrastructure for academia, it is key to have one national solution for all HEI, to directly fund the one national solution, which means direct funding from the state or other stakeholders, rather than via different HEI; and to make the national solution appear independent and neutral. They also recommended adapting the review process for accessing international supercomputers such as LUMI via CSCS to make it more agile. It was also stated that the CSCS was a driving force behind the establishment of the LUMI supercomputer and its consortium. The CSCS is also perceived as visionary and extremely well connected internationally as well as with the suppliers. This is seen as an absolute asset that has put Switzerland in a good position – one that must be maintained.

For these reasons, the Council concluded that, while it would continue to recommend the existing supercomputing provider, its governance should be adjusted in such a way that it fulfils and reflects its national mandate. By doing this, the Council aims to ensure that this important supercomputing centre is independent of the interests of any single institution and serves all ERI stakeholders in Switzerland equally, including those outside the ETH domain.

### For Tier 2: National high-performance computing

- To promote a national high-performance computing infrastructure which serves a wide variety of users who do not require supercomputing resources, and which integrates or builds on existing computing resources.

## Why?

National interviews revealed that, while Switzerland currently lacks a national HPC infrastructure to serve a broad range of users who do not require the most advanced computing resources (i.e. supercomputing), such a tier is necessary given the limited local or regional computing resources available to HEI. This is why the SSC recommends establishing such a tier. Organisations established in and for the Swiss higher education sector that act as ICT providers for HEI and are aware of digital sovereignty already exist, an example being Switch. Such organisations could scale up their current computing resources and services to provide the recommended national high-performance computing tier for a wide variety of users of HEI throughout Switzerland. Such an approach would create synergies.

<sup>149</sup> Access to LUMI: <https://www.lumi-supercomputer.eu/get-started-2021/users-in-switzerland/> and through the activities of Helvetic AI Resources, Technologies and Services (HEARTS), see <https://csc.fi/en/news/eurohpc-ju-se-lects-ai-factory-antennas-to-broaden-ai-factories-initiative/> and <https://goba.swiss/en/switzerland-joins-europe-an-ai-factory-network-with-new-hearts-antenna/> (last checked on 26.01.26).

### For Tier 3: Regional compute centres<sup>150</sup>

- To set up local or regional AI expert groups at higher education institutions that support users locally and direct them to the appropriate computing resources within the tiered system, as required.

## Why?

‘Regional compute centres’ refers to the central HPC services provided by many cantonal universities and UAS. National interviewees emphasised the importance of having local AI expertise that can provide qualified answers to various AI-related questions in the shortest possible time. Some international interviewees stated that, for a large computing system with a small number of users, it is essential to provide expertise in understanding such systems (e.g., supercomputers). Therefore, machine operation and user support cannot be separated. However, for smaller systems with many users, computing system operations and user support should be separate. Other international supercomputer centres employ staff solely to run and operate the supercomputer, as there are centrally funded expert communities that provide the necessary user support. The SSC therefore recommends that higher education institutions establish local or regional AI expert groups to serve as contact points, provide the necessary local expertise and advise users on how to access the computing capacities they need within the tier system.

- The tiers should provide computing resources that include all the necessary components, services and resources required to make computing power usable by end users, thereby also functioning as pools of national experts for user and scientific support including AI developments.

## Why?

For almost all of the national interviewees, computing resources include not only the hardware itself, but also all the necessary components and services required to make the computing power usable by end users. This includes hardware and software (e.g., to run and use the hardware for its intended purpose, such as academic research), personnel (e.g., to set up and maintain the hardware, code the necessary software, run it on the hardware and support end users, such as researchers, in using the computing infrastructure, as well as providing scientific support, such as scientific programming and analysis), and other resources, such as energy and water for cooling (see Appendix 4.3 for definitions). All international interviewees also stated that user support is crucial, as it plays a central role in enabling users, such as researchers, to make use of the computing resources. This is why the SSC recommends that all tiers provide computing resources as defined. The Council is convinced designing computing resources as such would simultaneously create a national pool of experts in user and scientific support. This is necessary in order to be able to respond to current and future developments and methodologies in the field of AI in higher education.

<sup>150</sup> Example: HEI: central HPC (on-premise solutions) of cantonal universities or of UAS.

## 2.3 Agile computing infrastructure strategy board

- To establish a national, independent and neutral strategic board that assesses and decides on the design of a future-oriented computing infrastructure in form of a tiered system. The strategic board is responsible for:
  - designing the tiered system strategically, based on its continuous assessment of national large-scale computing resource needs and international computing infrastructure developments, taking into account multilevel governance. The national strategy should be based on the guiding principles outlined above.
  - monitoring the implementation and performance of the national strategy by the single tiers of the tiered computing infrastructure system.

The Federal Council should appoint this strategic board as an independent commission of experts. It should be equipped with the necessary powers to fulfil its responsibilities, if necessary on a legal basis.

### Why?

As AI computing capacities are evolving rapidly and cloud computing is converging with HPC, it is impossible to make recommendations for the exact design or structure of future computing resources that will still be relevant in five years' time. This is why a strategic board must continually evaluate national large-scale computing requirements and international IT infrastructure developments to design a computing infrastructure for academic purposes in Switzerland. Since the aim of this strategic board is to design a strategy in line with the guiding principles already outlined for a computing infrastructure for Switzerland's diverse ERI landscape, it must be a national board that has decision power and is also independent, for example, of any single institution, and thus able to make neutral decisions. This guarantees serving ERI stakeholders throughout Switzerland, in-

cluding those beyond the technical disciplines. Likewise, the multi-governance system<sup>151</sup> of Swiss federalism should be considered, given that the ERI system is embedded within it. The Council emphasises that the recommended system (national strategy, tiered computing infrastructure system, strategic board) is intended to respond to rapid developments in the field. This means that agility must be an integral and practised part of the system. Therefore, the national strategy should not be rigid or prevent bottom-up initiatives. Rather, the focus is on 'systemness', i.e. the coordination of multiple components (i.e. stakeholders) that, when working together, create a network of activity that is more powerful than the actions of the individual parts alone. For this reason, too, the strategic board is responsible for monitoring the implementation and performance of the national strategy by the individual tiers of the computing infrastructure system. As this strategic board must be visionary, agile and competent in various areas (see bullet point below), the SSC believes that it should not be integrated into any of the existing organisations in the ERI domain. Instead, the Council believes that this rapidly evolving topic requires a new independent expert commission, appointed by the Federal Council and equipped with the decision-making power necessary to fulfil its responsibilities and tasks. If necessary, for example to regulate its decision-making powers, this expert commission, responsible for developing and implementing this national AI strategy for the ERI domain in accordance with the recommendations set out here, should be enshrined in law.

<sup>151</sup> Trein, J. P. Multilevel Governance and the Digital Transformation of Research Libraries: An Analytical Framework. Expert Report for the Swiss Science Council SSC. 2025. doi.org/10.5281/zenodo.17378458.

- The strategic board should comprise all relevant competencies required to undertake their accountabilities, including technical, scientific, strategic, legal and financial expertise, simultaneously representing users and interest groups of national and regional computing resources of higher education institutions, with the size of the board allowing for efficient steering.

strategic board should also be users and representatives of national computing resources and regional data centres. The size of the board must allow for efficient steering. The SSC envisages a competent and representative board that is lean and acts with agility and efficiency because it is equipped with the necessary decision-making power and considers the computing resource needs of the entire ERI system.

## Why?

To develop a national strategy for computing resources that enables the use of AI for ERI in Switzerland, the strategic board must bring together the necessary expertise to carry out its duties. This includes, for example, technical, scientific, strategic, legal and financial expertise. Technical expertise includes in-depth knowledge of advanced hardware architectures, accelerators, specialised chips, system software, programming, optimisation, performance tuning, network and data infrastructures, data management, and emerging technologies (e.g., quantum computing). Scientific expertise involves domain-specific research knowledge, such as understanding the unique computational, data and simulation requirements of different fields of research (e.g., climate modelling, genomics), to ensure that HPC system designs align with real-world scientific demands. Strategic expertise involves interdisciplinary collaboration and strategic foresight to foster collaboration among technical, scientific, policy and business experts while maintaining a long-term strategic vision that aligns computing infrastructure with evolving research and market trends. Legal expertise brings knowledge of supply chain dynamics within the ecosystem, including an understanding of supplier relationships, intellectual property considerations and compliance with regulatory frameworks to mitigate risks and ensure a legally sound computing infrastructure, including associated storage and archiving systems. This includes risk management and regulatory compliance, such as managing legal and cybersecurity risks and adhering to export controls, data privacy regulations, and other legal obligations to ensure strategic investments are resilient and compliant. Financial expertise involves strategic investment analysis and financial acumen. This involves conducting thorough cost-benefit analyses and managing long-term investment strategies while balancing performance, scalability and fiscal responsibility. The experts who possess the essential competencies required for the tasks of this

## 2.4 Financial implications

- To adapt funding in order to reimburse computing resources in their entirety, financing can only partially rely on grants, as these only provide for short-term costs; computing resources, however, require long-term commitment and investment. Therefore, adequate long-term funding must be provided for all tiers. In general, computing resources should be provided in a transparent, cost-effective manner, meaning that costs must be covered but no profit should be made out of higher education institutions.

### Why?

National interviewees stated that alternative funding mechanisms should be explored in order to support sustainable computing infrastructure. They do not consider funding computing resources via grant proposals or grants with matching funds from universities to be ideal for strategic, large-scale computing because such a system is neither predictable nor scalable in the long term. Most of the international supercomputer centres interviewed are publicly financed in full or offer their services free of charge to academia. Some stated that it is difficult to find the 'correct' pricing model for academia. The challenge lies in ensuring that costs are covered, while also remaining competitive and not undercutting industry prices. One international interviewee saw a risk in providing computing resources free of charge to academia, as some researchers might waste these resources while others with great ideas might not get the resources they need. In general, academia does not have high amounts of budget for computing or too small a budget for what AI research costs, and this budget often does not include user support. This is because it is easier to receive grant money for hardware, but user support funding is missing from this financing scheme, which presents a challenge. As outlined previously, computing resources include all the necessary components, services and resources required to make computing power available for end users (see Appendix 4.3 for definitions). This is why computing resources must be financed in their entirety, as such a budget is a prerequisite for the feasibility of a project. Therefore, the SSC concluded that, due to their cost, computing resources as defined here require long-term financial investment. Consequently, national fund-

ing needs to reimburse computing resources in full, as financing can only partially rely on grants, which only provide short-term investment. In general, computing resources should be provided in a cost-effective manner, meaning that costs must be covered, but no profit should be made out of HEI.

## 2.5 What are the benefits of implementing these recommendations?

The recommended system creates a modular computing infrastructure that serves different user needs in terms of computing resources, user support and scientific support across Switzerland. The process is steered, the infrastructure is modular and leverages synergies, helping to minimise duplication while allowing agility. This lays one of the foundation stones required for computationally intensive research across a wide range of disciplines.

### 3 What other foundation stones are needed?

Computing resources are not the only prerequisite for enabling the use of AI in the ERI domain. To fully leverage these computing resources, findable, accessible, interoperable and reusable data across all disciplines is required. This applies not only to research data, but also to primary data that has not been collected for research purposes. As outlined in the guiding principles of the recommended national AI strategy, a coherent national data strategy is therefore needed, as a siloed approach to computing resources and data management limits progress.<sup>152</sup> Research and provision must be brought together. Therefore, as sector-specific data spaces are established (e.g., in health or mobility), they must be designed to interconnect.<sup>153</sup> This means that, at some point, it must be possible to combine data from different spaces in order to address cross-domain questions, whether supply- or research-oriented. After all, it cannot be predicted which questions will be important to tackle future challenges and research needs, for example in the event of a future crisis. Consequently, connectivity and flexibility must be embedded in the design of these spaces from the outset. This is something that can be learned from hyperscalers. They effectively integrate data and computing resources to create scalable networks that foster innovation. Realising this vision may necessitate bold, transformative decisions and a coordinated national effort. For some stakeholders, this may entail shifting towards a more collaborative and cross-sectoral mindset.

152 Kolbe-Guyot M. and Finger M. Data strategy, policy and regulation for Switzerland. EPFL Centre for Digital Trust. C4DT Insight 2 May 2025. <https://drive.switch.ch/index.php/s/nswBpwOvFQJKFBX> (last checked on 26.01.26; report not available in English).

153 Federal Chancellery FCh. Digital Transformation and ICT Governance. Architectural vision and principles of the Swiss data ecosystem. 12 May 2025. [https://www.bk.admin.ch/bk/en/home/digitale-transformation-ikt-lenkung/datenoeokosystem\\_schweiz/grundlagen-fuer-datenraeume/architekturvisionprinzipiendatenoeokosystemsschweiz.html](https://www.bk.admin.ch/bk/en/home/digitale-transformation-ikt-lenkung/datenoeokosystem_schweiz/grundlagen-fuer-datenraeume/architekturvisionprinzipiendatenoeokosystemsschweiz.html) (last checked on 26.01.26; report not available in English).

## 4 Annex

### 4.1 National and international interview guides

#### 4.1.1 National interview guide for determining computing infrastructure resources

##### Definitions

Term	Definition
Cloud computing	is the paid use of computing resources that are provided by public (e.g., <a href="#">Swiss National Supercomputing Centre, CSCS</a> ) and private providers (e.g., Amazon, Microsoft) via the Internet, whereby large amounts of computing power can be accessed on demand, for example for artificial intelligence such as machine learning.
Computing infrastructure	are quantifiable amounts of in-house computing resources that can be requested, allocated and consumed for computing activities such as machine learning.
Computing resources	includes <ul style="list-style-type: none"> <li>• computing power (Central Processing Unit CPU in Millicores; Graphics Processing Unit GPU in floating point operations per second FLOPs),</li> <li>• memory (random access memory RAM in Bytes)</li> <li>• networks (client-server architecture, Peer-to-peer P2P architecture)</li> <li>• storage (direct area storage DAS, network-based storage (e.g., network-attached storage NAS, storage area network SAN))</li> </ul>
Machine Learning (ML)	is a subfield of artificial intelligence that develops and uses computer systems that are able to learn and adapt without following explicit instructions, by using algorithms and statistical models to analyse and draw inferences from patterns in data.

##### Context

Research using large quantities of data has become increasingly important in the last decade. Machine learning, which offers extensive opportunities in data-driven research, has certainly fuelled this trend. A prerequisite for data-intensive and therefore computing power-intensive analyses is sufficient computing infrastructure. The SSC is therefore investigating what computing infrastructure is currently available in academic research and to what extent cloud computing is used for academic research. The aim is also to assess the needs of the Swiss research community in terms of computing infrastructure and cloud computing over the next 5 to 10 years.

##### Questions

###### General questions concerning computing infrastructure resources:

1. We define computing infrastructure resources as shown in the table (see above). Do you agree with these definitions or would you categorise your computing infrastructure differently?
2. Which in-house, private and/or public cloud computing infrastructure resources do you use? What percentage of in-house, of private and of public cloud computing infrastructure resources do you use?

3. Why did you choose in-house or cloud computing infrastructure resources and are you satisfied with your choice?
4. What is your vision for computing infrastructure resources (in-house and cloud computing, from private or public providers) in the next 5 to 10 years?

**Computing infrastructure resources (in-house):**

5. Do you share your computing infrastructure with other entities? If yes, with which entities and why?
6. What are the key indicators (computing power, memory, networks, storage) of your computing infrastructure?
7. What were the running costs of your computing infrastructure per year in the last 5 years?
  - a. How many people (in full time-equivalent) are employed to run the in-house computing infrastructure, as a service for the users in your institution (researchers, students, etc.)?
  - b. Did you have to hire additional personnel because of increased computing needs by your users in the past 5 years?
8. What were the energy costs of your computing infrastructure per year in the last 5 years?
9. What is the utilisation rate of your computing infrastructure?
10. Do you plan to invest in your computing infrastructure in the coming 5 to 10 years?
  - a. If yes, to what extent? (in terms of computing power, memory, networks, storage, investment cost, running costs, estimated energy costs)
  - b. If yes, what drives this investment? (how much % is for ML, more general AI, 'classical' scientific computation including both theoretical modelling and experimental data reduction, applications-oriented work such as weather prediction etc.)
11. What are your back-up and disaster recovery plans, and what are the associated providers and costs?

**Cloud computing (by private providers or public providers (see question 2))**

12. What are the key indicators (computing power, memory, networks, storage) of the purchased cloud computing?
13. What were the running costs of the cloud computing per year in the last 5 years?
  - a. How many people (in full time-equivalent) are employed to run the private/public cloud computing infrastructure, as a service for the users in your institution (researchers, students, etc.)?
  - b. Did you have to hire additional personnel because of increased computing needs by your users in the past 5 years?
14. What were the energy costs of the cloud computing per year in the last 5 years?
15. To what extent do students and professors use free cloud computing options?
  - a. To what extent do professors receive and use cloud computing gifts (e.g., a fixed number of GPU hours they can use for free)? (see also question 18)
16. Which cloud computing provider(s) do you use?
  - a. Where is the cloud located (e.g., CH, EU, US, China)?
17. How are data protected? What security/encryption is used, and what is the cloud computing provider allowed to do (e.g., with the data) and what not? Can the provider be asked to give access to the data by any government or other organisation?
18. In what way is your contract with the cloud computing provider different from those of other higher education institutions?

19. Do you plan to extent your use of cloud computing from private and public providers in the coming 5 to 10 years?
- a. If yes, to what extent? (in terms of computing power, memory, networks, storage, investment cost, running costs, estimated energy costs)
  - b. If yes, what drives this investment? (how much % is for ML, more general AI, 'classical' scientific computation including both theoretical modelling and experimental data reduction, applications-oriented work such as weather prediction etc.)
  - c. Is data residency, sovereignty, access and protection a concern for you? If so, why?

**Additional questions:**

20. Can you answer these questions for your entire institution? Y/N (why not)
- a. Are there new types of services, applications and users for computing infrastructure resources (e.g., for digitisation of university libraries)?
21. How much effort does it take to answer the questions? Small/large (why large)
22. How accurate are the numbers you have provided? (Are they estimates or are they easily quantifiable?)
23. What were the difficulties in collecting these numbers?
24. Also of interest, but not concerning computing infrastructure, but AI talent:
- a. Does your institution have difficulties in retaining and/or attracting top AI talent in academic research (e.g., due to a brain drain to industry)? Y/N (why/why not)
  - b. Do you have plans to increase the pool of top AI talent? Y/N (why not?)

## 4.1.2 Interview guide for international supercomputer centres

### Definitions

Term	Definition
Cloud computing	is the paid use of computing resources that are provided by public (e.g., <u>Swiss National Supercomputing Centre, CSCS</u> , EuroHPC) and private providers (e. g. Amazon, Microsoft) via the Internet, whereby large amounts of computing power can be accessed on demand, for example for artificial intelligence such as machine learning.
Computing infrastructure	are quantifiable amounts of in-house computing resources that can be requested, allocated and consumed for computing activities such as machine learning.
Computing resources	<ul style="list-style-type: none"> <li>• includes hardware: <ul style="list-style-type: none"> <li>– computing power (Central Processing Unit CPU in Millicores; Graphics Processing Unit GPU in floating point operations per second FLOPs),</li> <li>– memory (random access memory RAM in Bytes)</li> <li>– networks (client-server architecture, Peer-to-peer P2P architecture)</li> <li>– storage (direct area storage DAS, network-based storage (e.g., network-attached storage NAS, storage area network SAN)</li> </ul> </li> <li>• includes software to run and use the hardware for its purpose (e.g., academic research)</li> <li>• includes personnel to set up the computing resources hardware and to maintain it, as well as personnel to code the software and run it on the hardware</li> <li>• includes other associated materials such as energy consumption</li> </ul>
Machine Learning (ML)	is a subfield of artificial intelligence that develops and uses computer systems that are able to learn and adapt without following explicit instructions, by using algorithms and statistical models to analyse and draw inferences from patterns in data.

### Context

Research using large quantities of data has become increasingly important in the last decade. Machine learning, which offers extensive opportunities in data-driven research, has certainly fuelled this trend. A prerequisite for data-intensive and therefore computing power-intensive analyses is sufficient computing infrastructure. The SSC is therefore investigating what computing infrastructure is currently available in academic research and to what extent cloud computing is used for academic research. The aim is also to assess the needs of the Swiss research community in terms of computing infrastructure and cloud computing over the next 5 to 10 years, and how it could network with and contribute to international (mainly European) supercomputer centres.

### Questions

1. We define computing infrastructure resources as shown in the table (see above). Do you agree with these definitions, or would you categorise your computing infrastructure differently?
2. What are the computing infrastructure resources of your supercomputing centre?
  - a. Do you provide in-house and/or public cloud computing and/or private cloud computing infrastructure resources? If yes to which percentages?

- b. What are the key indicators (computing power, memory, networks, storage) of your computing infrastructure resources?
  - c. How is access to the computing infrastructure resources regulated and who has access at which costs?
  - d. How is data transfer (upload, download) regulated and what are the associated costs?
  - e. How are data protected? What security/ encryption is used, and with whom is/can it be shared? (Can sensitive data (e.g., medical data) be calculated?)
  - f. What is the utilisation rate of your computing infrastructure resources?
  - g. What percentage of computing infrastructure resources are used for academic research and what percentage for industrial research?
  - h. What were the running costs of your computing infrastructure resources per year in the last 5 years?
  - i. What were the energy costs of your computing infrastructure resources per year in the last 5 years?
  - j. What are your back-up and disaster recovery plans, and what are the associated providers and costs?
  - k. Does your supercomputer centre have difficulties in retaining and/or attracting top AI talent in academic research (e.g., due to a brain drain to industry)?
3. What lessons did you learn when you set up your computing infrastructure resources and what would you do differently?
  4. How are the computing infrastructure resources financed and for how long are finances secured?
  5. Are you using public-private partnership for funding and what is the pro and con of this partnership?
  6. What is your vision for computing infrastructure resources (in-house and cloud computing, from private or public providers) in the next 5 to 10 years and what is your long-term roadmap for the computing infrastructure resources?
    - a. What are the current challenges and limitations for your vision to come true and which prerequisites must be met?
    - b. Do you also plan to use or are you already using alternative new technologies such as classical (CPU, GPU, etc.) and quantum computing? If yes, how and why?
    - c. Do you also plan to use or are you already using alternative hardware providers (others than Nvidia)? If yes, for which purposes and why?
    - d. Do you have access to all the technology (e.g., chips) you need and if not, how are you going to deal with this challenge?
  7. What are the possibilities for other countries to have a greater share in your computing infrastructure resources in the future, so that their researchers have (more) access to it?

### **4.1.3 Standardised questions for validation interviews**

1. Are the recommendations understandable?
2. Would you in general support the recommendations from your HEI's perspective? If so, why? If not, why not?
3. In your opinion, are there any aspects that are not included in the recommendations but should be, i.e. is there anything important missing? If so, what is it and why is it important?
4. In your opinion, are there any aspects that are misrepresented in the recommendations? If so, which ones and why do you think they are misrepresented?
5. In your opinion, are there any challenges in implementing these recommendations? If so, what are they and why?

## 4.2 Working process

The table below only lists meetings with the entire Council, meetings with the working group set up for this topic (consisting of Council members and a member of the SSC secretariat) and national and international interviews conducted for this project.

Date	Meeting or Interview	Purpose
07.05.2024	Plenary meeting	Decision to address the topic of AI in the SSC's working programme.
14.05.2024	Working Group Meeting AI	Discussion of topics that could be addressed in the field of AI that are relevant for ERI systems.
30.05.2024	Working Group Meeting AI	Discussion of topics that could be addressed in the field of AI that are relevant for the Swiss ERI system.
19.06.2024	Working Group Meeting AI	Decision to work on the topic of AI and computing infrastructure.
01.07.2024	Plenary meeting	Decision to address the topic of AI and computing infrastructure as it is relevant for the Swiss ERI system. Decision about the method how to tackle this topic.
28.08.2024	Working Group Meeting AI Infrastructure	Feedback on the national interview guide. Selection of the first three vice-presidents or (research) rectors to contact for the pilot study. Definition of the main reasons for the pilot study that should be mentioned in the letters to contact HEI.
23.09.2024	Plenary meeting	Approval of the Project Proposal AI and Infrastructure outlining the question to address, the method, the output and the timeline.
01.11.2024	1. Interview Swiss HEI (Pilot study)	Assessing computing infrastructure resources and future needs with national interview guide; validation of national interview guide.
13.11.2024	2. Interview Swiss UAS (Pilot study)	Assessing computing infrastructure resources and future needs with national interview guide; validation of national interview guide.
20.11.2024	3. Interview Swiss HEI (Pilot study)	Assessing computing infrastructure resources and future needs with national interview guide; validation of national interview guide.
26.11.2024	Plenary meeting	Discussion of the results of the pilot study and decision to conduct the entire set of interviews using the national interview guide.

16.12.2024	Working Group Meeting AI Infrastructure	Decision on which HEI and organisations to interview next. Decision on which international supercomputer centres to interview.
20.01.2025	4. Interview Swiss HEI	Assessing computing infrastructure resources and future needs with national interview guide.
29.01.2025	5. Interview Swiss HEI	Assessing computing infrastructure resources and future needs with national interview guide.
29.01.2025	Working Group Meeting AI Infrastructure	Discussion of lessons learned from the national interviews conducted so far. Discussion what questions to ask the international supercomputer centres.
03.02.2025	6. Interview Swiss UAS	Assessing computing infrastructure resources and future needs with national interview guide.
11.02.2025	7. Interview Swiss HEI	Assessing computing infrastructure resources and future needs with national interview guide.
24.02.2025	8. Interview Swiss organisation	Assessing computing infrastructure resources and future needs with national interview guide.
25.02.2025	Plenary meeting	Discussion on lessons learned from national interviews conducted so far. Discussion on suggested solutions for future computing resources.
05.03.2025	9. Interview Swiss HEI	Assessing computing infrastructure resources and future needs with national interview guide.
10.03.2025	10. Interview Swiss HEI	Assessing computing infrastructure resources and future needs with national interview guide.
02.04.2025	11. Interview Swiss organisation	Assessing computing infrastructure resources and future needs with national interview guide.
02.04.2025	12. Interview Swiss HEI institute	Assessing future computing infrastructure resource needs in preparation for the international interviews.
03.04.2025	Working Group Meeting AI Infrastructure	Discussion of lessons learned from the national interviews conducted so far. Feedback on international interview guide for international supercomputer centres.
03.04.2025	13. Interview Swiss HEI	Assessing computing infrastructure resources and future needs with national interview guide.
06.05.2025	Plenary meeting	Discussion on lessons learned from all national interviews conducted. Discussion on suggested solutions for future computing resources. Information on international interviews planned.
14.05.2025	1. Interview international supercomputing centre	Assessing computing infrastructure resources and future needs with international interview guide.

02.06.2025	Working Group Meeting AI Infrastructure	Discussion on how to develop recommendations based on the results of the national interviews. Information on the first international interview.
03.06.2025	2. Interview international supercomputing centre	Assessing computing infrastructure resources and future needs with international interview guide.
18.06.2025	3. Interview international supercomputing centre	Assessing computing infrastructure resources and future needs with international interview guide.
23.06.2025	4. Interview international supercomputing centre	Assessing computing infrastructure resources and future needs with international interview guide.
30.06.2025	Working Group Meeting AI Infrastructure	Discussion and decision on which issues and topics identified in the national and international interviews recommendations should be developed (part 1).
01.07.2025	Plenary meeting	Discussion on lessons learned from all international interviews conducted.
21.08.2025	Working Group Meeting AI Infrastructure	Discussion and decision on which issues and topics identified in the national and international interviews recommendations should be developed (part 2).
05.09.2025	Working Group Meeting AI Infrastructure	Feedback on the drafted recommendations based on results of the national and international interviews and the previous decision of the working group. Decision to share the recommendations with some national stakeholders to receive feedback.
11.–15. 09.2025	Working Group Meeting AI Infrastructure	Written feedback and approval of the executive summary and the recommendations for AI and computing infrastructure that are presented in the next plenary meeting.
23.09.2025	Plenary meeting	Approval executive summary and the recommendations for AI and computing infrastructure.
09.10.2025	Working Group Meeting AI Infrastructure	Discussion and decision on the national interviewees with whom the recommendations will be shared to receive feedback. Decision on the structure of the accompanying report and the timeline.
24.10.2025	1. Interview with Swiss UAS	To receive feedback on the recommendations for AI and computing infrastructure.
03.11.2025	2. Interview with Swiss organisation	To receive feedback on the recommendations for AI and computing infrastructure.
06.11.2025	Working Group Meeting AI Infrastructure	Discussion on received feedback from national stakeholders on the recommendations. Discussion if or how to adapt the recommendations according to the feedback received.

10.11.2025	3. Interview with Swiss HEI	To receive feedback on the recommendations for AI and computing infrastructure.
25.11.2025	Plenary meeting	Discussion on received feedback from national stakeholders on the recommendations for AI and computing infrastructure.
03.12.2025	4. Interview with Swiss HEI	To receive feedback on the recommendations for AI and computing infrastructure.
18.12.2025	5. Interview with Swiss HEI	To receive feedback on the recommendations for AI and computing infrastructure.
15.01.2026	Working Group Meeting AI Infrastructure	Decision if or how to adapt the recommendations according to the feedback received by the second round of national interviews conducted. Approval of draft AI and computing infrastructure report.
23.-24. 02.2026	Plenary meeting	Approval of the adapted recommendations for AI and computing infrastructure and of the AI and computing infrastructure report.

## 4.3 Definitions

Term	Definition
Cloud computing	is the use of computing resources either from a public cloud or a private cloud. Public cloud refers to a service owned and operated by third-party providers like AWS, Azure, Google. The cloud infrastructure of a public cloud is shared among multiple organisations but logically segregated. Private cloud refers to a dedicated infrastructure either hosted on-premises or by third-party providers. The private cloud is owned by one user only and is not shared with other organisations or customers.
Computing infrastructure	are quantifiable amounts of computing resources (in-house, private and/or public cloud) that can be requested, allocated and consumed for computing activities such as for example machine learning.
Computing resources	includes all the necessary components, services and resources required to make computing power usable for end users. For example: <ul style="list-style-type: none"> <li>• Hardware: <ul style="list-style-type: none"> <li>– computing power (Central Processing Unit CPU in Millicores; Graphics Processing Unit GPU in floating point operations per second FLOPs),</li> <li>– memory (random access memory RAM in Bytes)</li> <li>– networks (client-server architecture, Peer-to-peer P2P architecture, AI tailored network with extremely high bandwidth)</li> <li>– storage (direct area storage DAS, network-based storage (e.g., network-attached storage NAS, storage area network SAN, AI tailored storage systems)</li> </ul> </li> <li>• Software: to run and use the hardware for its purpose (e.g., academic research)</li> <li>• Personnel: to set up and maintain the computing resources hardware, to code the necessary software and to run it on the hardware, to support end users (e.g., researchers) in the use of the computing infrastructure as well as scientifically (e.g., scientific programming, analysis)</li> <li>• Other associated resources: like energy and water for cooling.</li> </ul>

## 4.4 Abbreviations

<b>Abbreviation</b>	<b>Signification</b>
AI	Artificial Intelligence
AIRR	AI Research Resource
CERN	Conseil Européen pour la Recherche Nucléaire
CHF	Swiss Franc(s)
CPU	Central Processing Unit
CSC	IT Center for Science Ltd. (Finland)
CSCS	Swiss National Supercomputing Centre
DCAI	Danish Centre for AI Innovation
DEP	Data Exploitation Platform
DSA	Data Science Africa
EIFO	Export and Investment Fund of Denmark
ELLIS	European Laboratory for Learning and Intelligent Systems
EOSC	European Open Science Cloud
EPFL	École Polytechnique Fédérale de Lausanne
ERI	Education, Research, Innovation
ETH	Eidgenössische Technische Hochschule
ETHZ	Eidgenössische Technische Hochschule Zürich
EuroHPC	European High Performance Computing
EuroHPC JU	European High Performance Computing Joint Undertaking
FAIR	Findable, Accessible, Interoperable, Reusable
FDFA	Federal Department of Foreign Affairs
FIN	Finland
GCS	Gauss Centre for Supercomputing
GÉANT	Gigabit European Academic Network
GPU	Graphics Processing Unit
HEARTS	HElvetic AI Resources, Technologies and Services
HEI	Higher Education Institution(s)
HLRS	High Performance Computing Center Stuttgart
HPC	High-Performance Computing
HPCN	High Performance Computing and Networking
ICAIN	International Computation and AI Network
ICT	Information and Communications Technology
IT	Information Technology
JSC	Jülich Supercomputing Centre
LLM	Large Language Model(s)
LRZ	Leibniz Supercomputing Centre, Garching
LUMI	Large Unified Modern Infrastructure; European supercomputer located in Finland
PRACE	Partnership for Advanced Computing in Europe
RAISE	Resource for AI Science in Europe
SDSC	Swiss Data Science Center
SME	Small and Medium-sized Enterprise(s)
SNAI	Swiss National AI Institute
SSC	Swiss Science Council
UAS	Universities of Applied Sciences

# Impressum

Schweizerischer Wissenschaftsrat SWR  
Conseil suisse de la science CSS  
Einsteinstrasse 2  
CH-3003 Bern  
T +41 (0)58 463 00 48  
[swr@swr.admin.ch](mailto:swr@swr.admin.ch)  
[www.wissenschaftsrat.ch](http://www.wissenschaftsrat.ch)

ISBN 978-3-906113-81-4  
DOI: [doi.org/10.5281/zenodo.18850058](https://doi.org/10.5281/zenodo.18850058)  
Bern 2026

Redaktion / Rédaction / Redazione:  
Geschäftsstelle SWR / Secrétariat CSS / Segreteria CSS

Lektorat / Lectorat / Revisione:  
Doris Tranter

Übersetzung / Traduction / Traduzione:  
Der Originaltext ist auf Englisch. Die Übersetzungen des «Management Summary» und der «Recommendations» auf Deutsch, Französisch und Italienisch basieren auf dem englischen Text.

Le texte original est en anglais. Les traductions du «Management Summary» et des «Recommendations» en allemand, français et italien sont basées sur le texte anglais.

Il testo originale è in inglese. Le traduzioni del «Management Summary» e delle «Recommendations» in tedesco, francese e italiano si basano sul testo inglese.

Konzept und Design / Concept et mise en page /  
Concetto e Design:  
Heyday Konzeption und Gestaltung GmbH

[LinkedIn:](#)



