

SciLifeLab AI Strategy

Ola Spjuth, AI Lead, SciLifeLab

2026-02-11

V1.0

Strategy reference group:

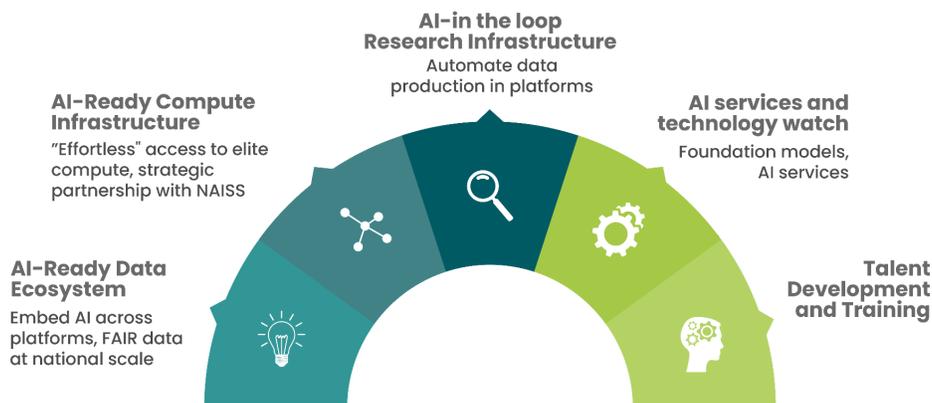
Magnus Boman, Ida-Maria Sintorn, Wei Ouyang, Patrick Bryant, Milka Doktorova, Jonas Hagberg, Andreas Luttnens

Summary

SciLifeLab is Sweden's national center for life science. Its infrastructure platforms generate vast amounts of data across genomics, proteomics, imaging, and chemical biology - exceeding 20 PB in 2025. Its research community includes over 30 new junior group leaders in the Data Driven Life Science (DDLs) Programme that drive the development of new tools and approaches to extract knowledge from such data. Together with strong national data management and bioinformatics services with close to 200 data science service staff and flagship data resources such as the Human Protein Atlas, SciLifeLab has a strong foundation for life science AI development and AI-driven discoveries. These capabilities, combined with SciLifeLab's collaborative role at the interface of academia, healthcare, and industry, represent a strategic advantage and a strong position to lead the development and integration of AI into Swedish and international life science.

The strategy is built on five interconnected pillars:

- **Pillar 1: AI-Ready Data Ecosystem:** Ensure SciLifeLab delivers FAIR, standardized, annotated, and interoperable datasets, directly available for machine learning, supported by initiatives like the Integrated Data Services (IDS) and flagship projects such as the Human Protein Atlas.
- **Pillar 2: AI-Ready Compute Infrastructure:** Provide reliable access to scalable compute and storage for life science AI development and application, including secure environments for sensitive data and hybrid solutions that integrate with industrial platforms.
- **Pillar 3: AI-in the loop Research Infrastructure:** Use AI to automate data production in platforms, identify data gaps and drive new technology development
- **Pillar 4: AI Tools & Technology Watch:** Deliver user-friendly AI tools such as foundation models, agents, and MLOps, while continuously monitoring global AI technology developments and translating them into life science tools.
- **Pillar 5: Talent Development & Training:** Expand AI training across all levels and roles, recruit international experts, and foster a connected and collaborative AI community within SciLifeLab.



Implementation of the AI Strategy will be anchored in SciLifeLab's data services, major research programs and partnerships. The DDLS program will emphasize life science AI in its next round of group leader recruitments and integrate AI as a central driver of research and training. Together with collaborations in computer science at SciLifeLab's partner universities, as well as WASP, this will ensure that new AI method developments can be linked with SciLifeLab's unique life science data generation capacity. New flagship projects such as AlphaCell will serve as drivers of innovation, integrating single cell multi-omics, imaging, and AI to build predictive, dynamic models of cellular systems.

Guiding principles include responsible and ethical AI aligned with GDPR and the EU AI Act, commitment to open science and FAIR data, agility in adopting new technologies, and long-term scalability and sustainability. Collaboration will extend nationally across universities, healthcare providers, other research infrastructures and funding agencies, and internationally through partnerships with EMBL, ELIXIR, EuroHPC, and others. Engagement with healthcare ensures translational and privacy-preserving AI models, while industry partnerships—from major tech companies to startups—provide pathways for innovation and commercialization.

Looking ahead, strategic areas for SciLifeLab include multimodal AI, domain-specific foundation models, reasoning agents, generative and simulation-based approaches, explainable AI, and automated labs. By acting now, SciLifeLab can establish Sweden as a global leader in responsible, impactful AI for life sciences and healthcare.

Contents

Summary	2
Contents	4
Introduction: AI in life science	5
SciLifeLab and AI	6
Vision	7
Strategic pillars	8
Pillar 1: AI-Ready Data Ecosystem	8
Pillar 2: AI-Ready Compute Infrastructure	9
Pillar 3: AI-in the loop Research Infrastructure	9
Pillar 4: AI services and technology watch	10
Pillar 5: Talent Development and Training	11
Strategic research programs	12
DDLS Program	12
WASP	12
Flagship Projects	13
Guiding Principles	13
Ethics, Responsible AI, and Regulatory Compliance	13
Open Science and FAIR Data Principles	13
Agility and Adaptability	14
Scalability and Sustainability	14
Collaboration, Partnerships and Innovation	14
National and International Collaborations	14
Healthcare	15
Industry	15
Innovation and Entrepreneurship	16
Outreach, Communication, and Public Engagement	16
Future Directions	17

Introduction: AI in life science

The field of life science is transforming from being largely observational to increasingly data-driven, a transition now massively accelerated by artificial intelligence (AI)¹. Groundbreaking achievements, such as DeepMind's AlphaFold, illustrate how AI can revolutionize fundamental biological understanding by accurately predicting protein structures on a proteome-wide scale—work that was awarded the Nobel Prize in Chemistry. More recently, advances in both specialized machine learning and generative foundation models have further accelerated the adoption of AI across the life sciences. But AI is not merely a technological upgrade - it is a paradigm shift that is fundamentally reshaping the way we explore, understand, and manipulate biology. It allows us to move from static, linear research processes toward highly adaptive, iterative cycles where experiments and computation continuously inform and accelerate each other in a closed feedback loop.

Life science poses unique challenges for AI. Nature has a very large diversity of biological systems, and biological data is large scale, extremely complex, heterogeneous, and noisy and often only allows indirect measurement of the physical and chemical nature of living systems. The “small data” problem, where the number of well-annotated examples is very limited, complicates model training. To be able to predict and re-engineer living systems, life science requires causal mechanistic understanding rather than simple correlation, which demands sophisticated models, careful interpretation and rigorous prediction testing. Ethical and privacy considerations are central when dealing with sensitive health or personal molecular data. These challenges mean that adopting AI in life science is not a matter of simply importing methods from other domains - it requires specific solutions.

Yet the opportunities are enormous. AI holds the promise to integrate and interpret the staggering complexity of biological systems, bringing together multi-omics data, imaging, and environmental measurements to form a holistic understanding of living systems. Generative AI can propose new hypotheses and guide researchers toward the most promising experiments or accelerate the research data processing and analyses. Combined with robotics and automation, AI has the potential to compress decades of research into a fraction of the time, turning discovery into a continuous and self-improving process.

Right now, we are in a global phase defined by massive investment in compute power and the development of large-scale foundation models, largely driven by Big Tech in the USA and China. Academia and smaller research organizations cannot compete on scale, but this is not where the greatest scientific breakthroughs necessarily occur. The most impactful discoveries often happen at the intersection of domains, where cutting-edge methods meet the right data and the most important questions. The next phase of AI in life science will depend not just on building ever-larger models but on coupling smart, targeted AI development for life science research with highest quality data generation and model driven experimental design.

For SciLifeLab, paving the way toward this future means building the infrastructure, expertise, and collaborations needed to fully harness the power of AI. This includes creating AI-ready datasets, ensuring researchers have access to “AI grade” computational resources, and embedding AI and automation into data production workflows. It also means fostering a culture where researchers are empowered to use AI responsibly and creatively. By taking these steps, SciLifeLab can help position Sweden at the forefront of the next wave of AI-enabled biological research and deliver discoveries that benefit science, healthcare, and society at large.

¹ In this document we use “AI” as the umbrella term used in the media and society, encompassing Machine Learning, Deep Learning, and related data-driven computational methods.

SciLifeLab and AI

SciLifeLab is Sweden's national research infrastructure for life science, bringing together universities, healthcare, and industry to enable world-class research and deliver societal impact. Its mission is to provide access to cutting-edge technologies, expertise, and data that empower researchers across disciplines. This makes SciLifeLab uniquely positioned to lead the integration of artificial intelligence into life science, turning AI into an important enabler of discovery.

SciLifeLab has significant strengths that form a strong foundation for AI. Its platforms have the capacity to generate data at national scale, covering genomics, proteomics, imaging, chemical biology, and more. During 2024, over 15 PB was generated at SciLifeLab, with expectations of 20+ PB generated in 2025. Sweden also benefits from a strong national bioinformatics infrastructure, giving researchers access to state-of-the-art computational tools and expertise. The SciLifeLab Data Centre has established the foundation for data services, and is contributing to storage, sharing, and management solutions that make data more accessible and reusable.

At the same time, important gaps must be addressed. Few life science research groups in Sweden are currently AI-native or design and carry out their research with AI as a central tool. Many platforms still do not deliver fully AI-ready data, lacking standardized metadata, harmonization, and interoperability needed for machine learning workflows. Platform staff and researchers also need significant upskilling to use AI effectively and responsibly. Addressing these challenges will be crucial to maximise the impact of AI in research.

An important step forward is the Integrated Data Services (IDS) initiative at SciLifeLab, launched in 2025. This initiative aims to ensure that platforms take greater responsibility for the quality and structure of the data they deliver to researchers and supports cross-platform data integration. The goal is to make data more discoverable, interoperable, and reusable—laying the groundwork for large-scale, AI-driven analyses. AI is a key driving force in this effort, motivating the development of harmonized, machine-actionable data for the next generation of integrative and predictive research.

However, Sweden already holds a wealth of valuable life science data that can power AI development. Flagship initiatives such as the Human Protein Atlas (HPA)² have created globally unique reference datasets on protein localization and expression across tissues and cell types. These resources represent a strategic advantage: when made AI-ready and integrated through initiatives like IDS, they enable the training of domain-specific foundation models and the creation of benchmark datasets that can drive methods development.

By building on these strengths, addressing weaknesses, and accelerating initiatives such as IDS, SciLifeLab will create the conditions for AI to be seamlessly embedded across its infrastructure. This will empower researchers to make faster, more reproducible discoveries and translate them into real benefits for science, healthcare, and society.

² <https://www.proteinatlas.org/>

Vision

SciLifeLab envisions a future where AI is a natural part of life science research, helping to turn data into insight and accelerating the journey to new discovery and real-world impact. AI will allow researchers to accelerate several steps of the research process and explore biological data at unprecedented scale and assist with generating new hypotheses and designing experiments. We fully expect this to lead to breakthroughs that improve health, protect the environment, and expand our understanding of life.

By leveraging SciLifeLab's unique strengths — primarily large-scale data generation in cutting-edge platforms — together with the collaborative and interdisciplinary spirit of Swedish life science, we will create an ecosystem where AI becomes a driver of innovation. Open, interoperable, and secure data resources, powerful computational resources, user-friendly data services, and targeted training will empower researchers to focus on discovery rather than engineering and e-infrastructure. With these efforts, SciLifeLab will not only lead nationally but also shape international standards for AI in life science, establishing themselves as global leaders in responsible, impactful, and collaborative AI-driven research.

Strategic pillars

SciLifeLab's AI strategy is built on five interconnected pillars that enable AI to become an important enabler of life science research. Together, they will unlock interoperable AI-ready data at SciLifeLab, ensure access to the computational power needed for tomorrow's models, rapidly translate emerging technologies into practical tools, and prepare researchers in Sweden to lead the next generation of AI-driven discovery.

Pillar 1: AI-Ready Data Ecosystem

AI fundamentally depends on data, and an AI-ready data ecosystem is crucial for unlocking the full potential of AI in life sciences. It requires datasets that are structured, standardized, and richly annotated to enable effective use in machine learning workflows. This includes consistent formats and directory structures, clear metadata, documented provenance, and appropriate labeling to support model training, validation, and interpretation. This is aligned with the FAIR (Findable, Accessible, Interoperable, Reusable) principles³ which in addition recommend that data must be easily discovered, interoperable so that it can be integrated across platforms and reused in diverse research contexts.

Building such an ecosystem is not simply a technical challenge but an organizational one, requiring coordinated investments across SciLifeLab's platforms and a shared commitment to making data a first-class research output. SciLifeLab needs to contribute to and lead development of common standards, harmonized metadata, and provide access to secure, scalable storage. Automated preprocessing and annotation will ensure data quality at scale and facilitate cross-platform, multi-modal analyses. Dedicated engineering capacity and technical expertise will be required to maintain and evolve these systems, and the IDS initiative at SciLifeLab constitutes an important initiative to this end.

Initial Priorities

- **Embed AI across SciLifeLab's platforms**, with upskilled personnel and automation to increase efficiency of core activities and user interactions, leveraging the IDS initiative for cross-platform integration.
- **Strengthen SciLifeLab platforms to deliver AI-ready data.** Drive technical development towards AI-readiness through automated workflows and computational infrastructure that supports efficient data access and integration. Data quality should be of highest priority, and ambition should be towards achieving machine-actionable FAIR data across all Swedish life science allowing researchers to work higher up in the data stack.
- **Build a portal of high-quality AI-ready datasets** to enable integrated analysis and AI methods development. Researchers affiliated with SciLifeLab and data-producing platforms could collaborate with this portal in assembling and publishing high-quality multi-omics/modal datasets. Work with scientists and platforms to publish high-quality multi-omics datasets. Leverage existing Swedish flagship projects such as HPA. Make data available via APIs.
- **Engage with the community and international standards work** to understand researcher needs, ensure SciLifeLab data is aligned with global best practices for interoperability and reuse, and that SciLifeLab contributes to the emerging community and international standards.

³ <https://www.nature.com/articles/sdata201618>

Pillar 2: AI-Ready Compute Infrastructure

Advanced AI applications in life science depend not only on high-quality data but also on powerful, reliable, and accessible computational infrastructure. As AI models grow in complexity and scale—particularly with the increasing use of deep learning and large foundation models—their demands for specialized hardware, such as GPUs, high-bandwidth interconnects, and fast data pipelines, have grown dramatically. At SciLifeLab, researchers across multiple domains are becoming increasingly reliant on computational resources for model training, large-scale inference, and simulation-driven experimentation. Without adequate e-infrastructure, these activities risk becoming bottlenecks that slow scientific progress.

Sweden's national high-performance computing resources, provided by NAISS, form the backbone of the country's research e-infrastructure, and recent investments such as the Arrhenius supercomputer and the creation of a Swedish AI Factory (Mimer) are promising steps toward meeting future needs. The current AI focus has been primarily on large-scale training workloads, with less support for inference, interactive workflows. Further, the storage needs for life science applications is growing continuously and is currently an area where further investments are needed. This gap particularly affects life science researchers who require secure, isolated environments for handling sensitive biomedical or clinical data, as well as flexible compute resources for exploratory and iterative work.

Initial Priorities

- **Establish a strong and strategic partnership with NAISS and other key infrastructure providers** to ensure access to compute and storage resources and influence the future HPC landscape in Sweden.
- **Invest in complementary computational resources**, secure environments, and user-friendly services that support AI development and deployment, as a complement to NAISS capabilities.
- **Adopt cloud technologies** where appropriate to enable interoperability with industrial AI services and frameworks.
- **Build central competence and expertise** to guide national-level decision-making on AI infrastructure and help researchers make efficient use of resources. Harmonize with the international community.
- **Establish close coordination and collaborations with university IT and information security functions** to ensure strategic alignment with our partner universities and their relevant efforts.

Pillar 3: AI-in the loop Research Infrastructure

AI-in-the-loop transforms SciLifeLab's platforms from static workflows into adaptive systems where experiments, analysis, and decision-making form a continuous feedback cycle. By integrating both specialized and foundational ML models directly with instruments, robotics, and data pipelines, infrastructure platforms can achieve higher data value and higher throughput. AI in the loop allows for example to: quickly optimize experimental parameters, perform data quality control in real time, intelligently adapt workflows based on the data already generated and avoid overproducing data not needed to reach a conclusive analysis, but instead automatically suggest follow-up experiments that maximize scientific value. This approach accelerates data production, improves reproducibility, and

enables more efficient use of resources. It is especially impactful for complex and multimodal workflows in e.g. imaging, genomics, structural and chemical biology, where AI can guide sample selection, adjust acquisition settings, integrate data across the workflow or highlight missing data needed to strengthen models.

Initial Priorities

- **Integrate AI into platform workflows** to enable adaptive data generation, real-time quality control, and dynamic adjustment of experimental parameters during data acquisition.
- **Build APIs and automation layers** that provide standardized, interoperable interfaces connecting instruments, robotics, data pipelines, and AI services across platforms.
- **Use AI to identify data gaps** by continuously analyzing incoming data and guiding targeted experiments that maximize information gain.
- **Pilot AI-driven automation projects** in high-impact areas such as multi-modal analyses to demonstrate scalable, reproducible, and accelerated research workflows.
- **Explore and develop virtual and co-scientist capabilities** to support researchers and platform staff through AI-assisted experimental design, protocol optimization, data interpretation, and decision-making.

Pillar 4: AI services and technology watch

Researchers across life science increasingly need accessible, reliable AI services that allow them to analyze complex datasets, build models, and automate workflows without managing e-infrastructure, while ensuring secure handling and reproducible pipelines for sensitive data. To meet these needs, SciLifeLab must provide centrally operated AI services tailored for life science, empowering scientists to focus on data interpretation and discovery rather than e-infrastructure. At the same time, because AI evolves rapidly, SciLifeLab should act as a bridge between global technological advances and their application in life sciences through a strong technology watch that scans emerging areas such as multimodal learning and agentic systems, assesses their relevance, and translates them into suitable tools and workflows. Prototyping and validation pipelines must support this process, testing promising technologies on local datasets and integrating them into services once they meet standards for quality, security, and interpretability. By combining robust shared services with continuous technology translation, SciLifeLab can ensure researchers have timely access to state-of-the-art AI capabilities, accelerating discovery and keeping the focus on answering biological questions.

Initial Priorities

- **Provide access to foundation models and LLM services** suited for biological and biomedical data, including locally hosted models.
- **Pilot and invest in agentic AI approaches** to support complex, multi-step scientific workflows and automation of life science experiments.
- **Run shared AI services** such as MLOps for reproducible deployment, monitoring, and maintenance of AI models in research contexts.

- **Establish a dedicated Tech Watch team** responsible for continuously monitoring global AI developments, assessing the relevance in life science, and keeping the SciLifeLab community informed about these.
- **Establish a community of practice** to coordinate AI-related activities, interact with WASP, explore and evaluate new approaches to translation from research to application, and support the development of new AI tools.

Pillar 5: Talent Development and Training

To fully realize the potential of AI in life sciences, SciLifeLab must invest in developing the people who will build, apply, and interpret AI methods across its platforms and technologies, including infrastructure providers and infrastructure users. This requires a multi-layered approach to training, starting with foundational AI literacy for researchers new to the field, continuing with advanced workshops for computational scientists and platform staff, and culminating in hands-on support for research groups actively applying AI in their projects. Training efforts should be closely aligned with SciLifeLab's tools, data, and services, ensuring that participants can directly apply what they learn. Community-building activities such as seminars, meetups, and hackathons can complement formal training and foster a collaborative, knowledge-sharing culture.

In parallel, SciLifeLab must focus on recruiting, developing, and retaining AI talent—including data engineers, machine learning specialists, research software engineers, and researchers working at the interface of AI and life science. These professionals are critical not only for building infrastructure and services but also for translating AI advances into meaningful scientific impact. Clear career paths, mentorship opportunities, and attractive packages for international experts will be key to strengthening the workforce.

Initial Priorities

- **Map current use and competence in AI** among SciLifeLab group leaders and fellows and use it to guide efforts towards talent development and training.
- **Expand AI-focused training at SciLifeLab**, enabled through the Training Hub, acknowledging the differences between model training and model use/inference. Provide tailored offerings for beginners, advanced users, and platform staff. Where relevant, facilitate discovery and access to international training opportunities, especially those offered by EMBO/EMBL. Focus on distinct needs of working with life science data (infrastructure providers, users of infrastructure, participants in SciLifeLab programs, e.g. DDLS Research School and Fellow programs), and on complementing training not available at Swedish universities.
- **Foster a connected and collaborative AI community** within SciLifeLab through cross-platform knowledge-sharing and collaboration. Organize hackathons, workshops, and peer-learning events to encourage gaining practical, hands-on experience, preferably on SciLifeLab datasets.
- **Actively recruit experts and provide mentorship** and career development opportunities to build long-term competence in AI. Foster joint mentoring of PhD students and postdocs. Work to eliminate administrative and organizational challenges to form interdisciplinary research groups in life science.

Strategic research programs

Strategic research programs are essential for advancing SciLifeLab's AI agenda, ensuring that resources, infrastructure, and collaborations are aligned with high-impact, long-term scientific goals. These programs should build on existing successful instruments while also launching bold new initiatives that reflect the rapidly changing AI and life science landscape.

DDLS Program

The Data-Driven Life Science (DDLS) program, funded by the Knut and Alice Wallenberg Foundation, is a national initiative hosted at SciLifeLab and designed to advance data-intensive research and training in areas such as cell and molecular biology, evolution and biodiversity, precision medicine, and epidemiology. There is an opportunity to adjust the DDLS program to strongly reflect the transformative role of AI in life sciences, positioning AI not just as a supporting tool but as a driver of new research questions, methodologies, and discoveries. Embedding AI in the program's strategic vision will ensure that its investments in data, infrastructure, and talent are future-proof and internationally competitive.

SciLifeLab should evolve DDLS to explicitly integrate AI as a central pillar of its research and training agenda. A central aspect is the recruitment of new research group leaders in life science with an AI-native/AI-first profile.

WASP

The Wallenberg AI, Autonomous Systems and Software Program (WASP), also funded by the Knut and Alice Wallenberg Foundation, is Sweden's largest research initiative in artificial intelligence, autonomous systems, and software development. Building collaborations between WASP and SciLifeLab offers strong potential to connect world-leading AI method development with impactful applications in life science. However it is challenging to form successful collaborations that are at the scientific frontier in both AI and life science. Here it is important to take advantage of existing data and grand challenges in life science, as well as SciLifeLab's capacity to generate new data to drive theoretical AI development.

There have been dedicated DDLS-WASP PhD projects funded, and larger NEST programs including researchers from both WASP and DDLS are about to start. These projects will foster interdisciplinary collaboration, train the next generation of AI-life science researchers, and create shared capabilities.

SciLifeLab should ensure the continuation of WASP-DDLS projects and, where possible, expand their scope and scale to include more research groups and thematic areas. By capitalizing on the capacity to generate novel large, multi-modal datasets - including benchmark datasets - SciLifeLab could steer and empower methods development within WASP.

Flagship Projects

Flagship projects are high-profile, large-scale initiatives that combine cutting-edge science, infrastructure, and collaboration. AlphaCell—a program to build a virtual human cell—is a prime example, integrating multi-omics, imaging, modeling, and AI to create predictive, dynamic models of cellular systems. In the program, AlphaCell will leverage the Human Protein Atlas as a rich reference for protein expression across tissues and cell types. This and similar initiatives act as national and international beacons of innovation.

SciLifeLab should prioritize the development and support of flagship AI-enabled projects, ensuring they have long-term funding, access to infrastructure, and international visibility.

Guiding Principles

Ethics, Responsible AI, and Regulatory Compliance

As AI becomes a core component of life science research, SciLifeLab must ensure that its development and use follow the highest standards of ethics and responsibility. AI systems applied to health, genomics, and other sensitive domains introduce risks related to bias, privacy, explainability, and accountability that must be carefully managed. Responsible AI at SciLifeLab means designing models that are transparent and interpretable, subjecting them to risk assessment and bias detection, and embedding ethical reflection throughout the entire lifecycle—from data collection and model training to deployment and evaluation. Regulatory compliance is a crucial part of this effort, requiring alignment with GDPR, the EU AI Act, and national guidelines. Working closely with legal and ethical experts, SciLifeLab will enable researchers to use AI tools with confidence, knowing that their work is compliant, fair, and trustworthy.

Ethical and compliant AI will be a hallmark of SciLifeLab's contribution to trustworthy, human-centered innovation.

Open Science and FAIR Data Principles

AI thrives on data, and its power is directly tied to the accessibility, quality, and reusability of that data. SciLifeLab's commitment to Open Science and FAIR principles provides the foundation for effective AI development and application. By making datasets, models, and workflows as open as possible—while respecting privacy and legal constraints—SciLifeLab can increase reproducibility, promote collaboration, and maximize the return on public investment in research. FAIR-aligned metadata, provenance tracking, and standardized formats make datasets machine-actionable and integrable across studies, enabling robust model training and meaningful benchmarking. Embedding these principles into every stage of research planning and execution will create an environment where AI can reach its full potential.

Open, FAIR, and machine-actionable data will fuel more powerful and generalizable AI models across life science.

Agility and Adaptability

The pace of AI development is extraordinary, and SciLifeLab must remain ready to respond to new breakthroughs. Modular infrastructure and flexible governance are essential so that new models, algorithms, and tools can be rapidly adopted without disrupting ongoing research. By fostering a culture of iteration and experimentation, SciLifeLab will support rapid prototyping of new AI workflows, evaluation of emerging techniques, and integration of successful approaches into production services. Organizational adaptability will also be critical, allowing technical staff and expertise to be redirected quickly toward high-impact opportunities.

Flexible systems and rapid iteration will keep SciLifeLab at the forefront of AI-enabled discovery.

Scalability and Sustainability

AI in life science will continue to grow in scale, with larger models, more data, and more users demanding compute and storage resources. SciLifeLab's approach must anticipate this growth, building scalable, shared services that can handle increasing demand without compromising performance or security. Automation of deployment, monitoring, and model versioning will be essential to maintain quality as complexity increases. Sustainability also requires stable funding, long-term staffing, and robust operational processes to avoid technical debt and ensure that services remain reliable and secure.

Long-term planning, stable teams, and continuous improvement will ensure that SciLifeLab's AI efforts remain impactful and relevant well into the future.

Collaboration, Partnerships and Innovation

Collaboration is essential for SciLifeLab's AI strategy to succeed. No single institution can address the full spectrum of challenges and opportunities presented by AI in life science. By fostering strong partnerships across sectors—nationally and internationally—SciLifeLab can leverage complementary strengths, accelerate knowledge exchange, and ensure real-world impact. These collaborations are also critical for ensuring that AI infrastructure, methods, and services are aligned with broader scientific, clinical, and technological ecosystems.

National and International Collaborations

As a national research infrastructure, SciLifeLab plays a central role in coordinating and enabling collaborations across Sweden's universities, healthcare providers, funding agencies, and innovation ecosystems. Strong national collaboration ensures that efforts are aligned, resources are shared efficiently, and expertise is pooled to address complex challenges in data, AI, and life science.

At the international level—primarily within Europe—SciLifeLab has the opportunity to both contribute to and benefit from a growing number of AI initiatives, research networks, and infrastructure programs. Collaborations with major infrastructures and organizations are essential to access large-scale computational resources, ensure data interoperability, and connect to cutting-edge AI research. By

strengthening these links, SciLifeLab can increase its visibility, secure resources, and position Sweden as a driving force in AI-enabled life science.

Priority collaborations and networks:

- **EuroHPC & AI Factories** – access to large-scale compute resources
- **EMBL, VIB, CSC** – links to leading European research infrastructures
- **EOSC (European Open Science Cloud)** – interoperability and FAIR data sharing
- **ELIXIR** – integration of life science data resources across Europe
- **EU research calls (Horizon Europe, Digital Europe)** – funding, networks, and leadership opportunities

SciLifeLab should actively engage in these networks, support researchers in joining and leading European projects, and strive to become not only a participant but a driver in shaping the future of AI in life sciences.

Healthcare

Healthcare systems are critical partners for enabling AI-driven research and translation. They provide access to unique and sensitive data—such as health records, genomics, imaging, and clinical outcomes—which are essential for developing and validating clinically relevant AI models. Collaborations with healthcare providers create the foundation for real-world testing, feedback, and deployment of AI tools, while ensuring alignment with ethical, legal, and regulatory standards. These partnerships also support the development of privacy-preserving and secure infrastructures that respect patient rights while enabling data-driven discovery and personalized medicine.

SciLifeLab should strengthen partnerships with healthcare providers to enable secure access to clinical data, co-develop translational AI models, and support ethical and impactful research.

Industry

Industry partnerships—ranging from major global technology firms, pharmaceutical and biotechnology companies to agile local startups—are increasingly important for driving AI innovation and scaling research outcomes. Collaborations with large technology companies offer access to advanced AI models, services, and compute resources, enabling SciLifeLab to stay at the forefront of technological development. At the same time, engaging with startups and biotech companies supports the development of specialized tools, rapid prototyping, and pathways for commercialization. These partnerships help bridge the gap between academic discovery and societal application, strengthening Sweden's position in the global innovation landscape.

SciLifeLab should build strategic partnerships with both large tech companies and startups to accelerate innovation, access advanced tools, and translate research into real-world applications.

Innovation and Entrepreneurship

SciLifeLab's AI methods and models have clear potential to become innovations for healthcare, industry, and society. A unique strength is the combination of rich data resources and the capacity to validate AI models across platforms. Opportunities include domain-specific models for diagnostics and drug discovery, curated benchmark datasets, and validation pipelines that ensure robustness and translational relevance. To realize this potential, SciLifeLab should link research outputs to Sweden's innovation ecosystem through IP support, spin-out facilitation, and partnerships with incubators, industry, and venture capital.

SciLifeLab should support the translation of AI research, leveraging its data and validation capacity to create impactful tools, companies, and societal value.

Outreach, Communication, and Public Engagement

As AI becomes a more prominent part of life science research and healthcare, clear and proactive communication is essential for building public trust, ensuring transparency, and fostering broad societal dialogue. SciLifeLab has a responsibility to engage not only with the scientific community, but also with policymakers, the public, and other stakeholders to explain how AI is being used, what benefits it offers, and how ethical and responsible practices should be ensured.

Outreach and communication efforts should aim to demystify AI, make research more accessible, and highlight the real-world impact of AI-driven discoveries in areas such as medicine, sustainability, and biodiversity. This includes publishing results in open-access formats, producing educational materials, and using diverse communication channels—including media, social platforms, and public events—to reach a wide audience. Proactive engagement is especially important in sensitive areas like genomics, patient data, and automated decision-making, where concerns about privacy, fairness, and transparency are high.

Public engagement also means listening. SciLifeLab should create opportunities for dialogue with patients, citizens, and civil society organizations to understand public values, expectations, and concerns around the use of AI in science and health. These insights are vital for shaping research agendas, guiding ethical frameworks, and ensuring that AI developments are aligned with societal needs and values.

SciLifeLab should communicate openly and accessibly about its use of AI, engage with the public and stakeholders to build trust, and ensure that public perspectives inform the direction of AI-driven research.

Future Directions

AI represents a paradigm shift in how science is conducted—transforming the pace, scale, and nature of discovery. This shift is already underway globally, with rapid developments in foundation models, autonomous experimentation, and AI-enabled scientific reasoning. For SciLifeLab, this moment presents a unique and time-sensitive opportunity to lead in shaping how these technologies are applied to life sciences and health.

The next generation of research infrastructure will not be defined solely by instruments and data, but by intelligent systems capable of learning, adapting, and accelerating every part of the scientific process. This includes virtual cells, self-driving labs, AI agents, and multimodal foundation models—all of which are evolving rapidly outside of Sweden. To remain competitive and relevant, SciLifeLab must act now to build the services, competencies, and collaborations needed to integrate these technologies into its ecosystem.

Examples of types of AI needed in future life science research

1. Multimodal AI

Life science data are highly heterogeneous—ranging from genomics and proteomics to imaging, clinical records, and environmental datasets. Future research will require AI systems that integrate multiple data modalities simultaneously, uncovering patterns invisible when each dataset is analyzed in isolation.

2. Foundation Models Specialized for Biology

While general-purpose foundation models are powerful, the life sciences will benefit from domain-adapted models trained on biological sequences, molecular structures, imaging, and omics data. These will enable applications such as protein design, drug discovery, and disease modeling at scale.

3. Reasoning AI and AI Agents

Next-generation AI will not just analyze data but also reason about hypotheses, design experiments, and orchestrate multi-step workflows. In life sciences, this means AI agents capable of planning experiments, querying databases, and interacting with robotic labs—leading toward self-driving laboratories and “robot scientists.”

4. Generative and Simulation-Based AI

The ability to generate realistic hypotheses, molecular structures, or even virtual cells and tissues will transform biology. Generative AI combined with mechanistic modeling can simulate biological processes, predict drug responses, and guide targeted experimentation.

5. Trustworthy and Explainable AI

Because decisions in health and biology are high-stakes, AI must be transparent, interpretable, and trustworthy. Future systems will include explainability, uncertainty quantification, and bias detection as core features, ensuring models are not only accurate but also scientifically credible and ethically sound.

6. Federated and Privacy-Preserving AI

Biomedical research often requires sensitive data. AI methods such as federated learning and synthetic data generation can enable collaboration across institutions and countries while respecting privacy and regulations.