

AI AND SDLABS

FOR SCIENTIFIC DISCOVERY &
TECH-TRANSFER

PROF. DR. MICHAEL J. BOJDYS

m.j.bojdys.02@cantab.net



EXECUTIVE SUMMARY

Self-driving laboratories (SDLabs) represent a transformative approach to research and development (R&D), combining artificial intelligence (AI), robotics, and digital tools to revolutionize workflows in deep-tech industries (e.g. materials, chemistry, biotechnology), and beyond. By automating experiments and decision-making processes, SDLabs enable faster and more efficient scientific discovery, reduce costs, and enhance reproducibility. This white paper explores SDLabs' potential, from accelerating innovation to addressing sustainability challenges, while highlighting critical ethical, cultural, and technological considerations. Furthermore, it examines the role of AI in technology transfer, showcasing success stories and actionable strategies for leveraging these advancements across sectors.

TABLE OF CONTENTS

Executive Summary2

Table of contents3

1: Introduction: Challenges in R&D and Tech-Transfer.....4

2: The Role of AI and SDLabs in Accelerating R&D6

 2.1: Defining SDLabs6

 2.2: AI-Driven Decision-Making7

 2.3: Benefits Over Traditional R&D Methods7

 2.4: Real-World Applications.....8

 2.5: Future Outlook8

3: Self-driving Tech-Transfer from Academia to Start-Ups9

 3.1: Overcoming Organizational Barriers9

 3.2: Fostering an Entrepreneurial Culture9

 3.3: The Role of SDLabs and AI in Tech-Transfer10

 3.4: Case Studies10

 3.5: Path Forward11

4: Contract Research 4.012

 4.1: Evolving CRO Dynamics.....12

 4.2: Regional Spotlights12

 4.3: "The Innovation Amplification Loop"13

5: Open & FAIR Science: Enabling Collaboration and Innovation.....15

 5.1: The Case for Open Science.....15

 5.2: Current Trends in Open & FAIR Science15

 5.3: The Future of Collaboration16

 5.4: Future Outlook: Emerging Models and Infrastructure17

Author.....19

Disclaimer.....19

1: INTRODUCTION: CHALLENGES IN R&D AND TECH-TRANSFER

Research and development (R&D) face a confluence of challenges as industries aim to accelerate innovation while aligning with sustainability goals. Traditional R&D methodologies, often reliant on manual experimentation and siloed knowledge, are struggling to keep pace with the complexity of modern scientific inquiries. The integration of advanced technologies, such as artificial intelligence (AI) and self-driving laboratories (SDLabs), offers a transformative solution to these bottlenecks, but their adoption is not without hurdles.

Challenges in R&D

1. **Time and Cost Constraints:** Traditional R&D processes are resource-intensive, requiring significant time and funding to conduct experiments, analyse data, and iterate findings. These lengthy cycles delay the introduction of new products and technologies, particularly in fields like pharmaceuticals, biotechnology, and materials science.¹
2. **Reproducibility Crisis:** Inconsistent experimental results have long plagued scientific discovery, undermining confidence in research outcomes. This issue arises from variability in methodologies, human error, and inadequate documentation.²
3. **Sustainability Pressures:** Industries are increasingly required to address environmental challenges by developing greener processes and products. However, traditional experimental setups often generate significant waste and rely on non-renewable resources.³
4. **Workforce Shifts:** With up to 30% of the workforce in sectors like chemistry projected to retire by 2030, industries face an impending knowledge gap that risks slowing innovation and progress.^{4,5}

Challenges in Tech Transfer

1. **Knowledge Silos:** Academic research often remains isolated, with limited mechanisms to translate discoveries into commercial products. This gap stifles the impact of cutting-edge research and slows technological adoption in industry.^{6,7}

¹ von Windheim, J.; Myers, B. "A lab-to-market roadmap for early-stage entrepreneurship" 2014. DOI: [10.1088/2053-1613/1/1/016001](https://doi.org/10.1088/2053-1613/1/1/016001)

² Miyakawa, T. "No raw data, no science: another possible source of the reproducibility crisis" 2020. DOI: [10.1186/s13041-020-0552-2](https://doi.org/10.1186/s13041-020-0552-2)

³ Deloitte "Sustainability regulation outlook 2024" 2024. [Link](#)

⁴ Royal Society of Chemistry (RSC) "The Future Chemistry Workforce and Educational Pathways" 2023. [Link](#)

⁵ Deloitte "The future of work in chemicals" 2021. [Link](#)

⁶ Apiou-Sbirlea, Gabriela, et al. "Anatomy and physiology of translation: the academic research imperative." 2015. DOI: [10.4155/cli.15.46](https://doi.org/10.4155/cli.15.46)

⁷ Hasenauer, J., et al. "From Planning Stage Towards FAIR Data: A Practical Metadatasheet For Biomedical Scientists" *Scientific Data*. 2024. DOI: [10.1038/s41597-024-03349-2](https://doi.org/10.1038/s41597-024-03349-2)

2. **Organizational Barriers:** Universities and research institutions are typically bound by bureaucratic processes, making it difficult to adopt agile methodologies and pilot innovative solutions.⁸
3. **Lack of Funding and Resources:** Early-stage innovations require substantial investment to move from proof-of-concept to scalable solutions, a hurdle many startups and academic spin-offs struggle to overcome.^{9,10}
4. **Ethical and Legal Complexities:** Intellectual property rights, data ownership, and ethical concerns create additional layers of complexity in transferring technology from lab to market.^{11,12,13}

The Role of SDLabs and AI Self-driving laboratories (SDLabs), powered by AI, are poised to address these challenges by automating experimentation, optimizing workflows, and fostering collaboration. SDLabs combine robotic platforms, real-time data analysis, and AI-driven decision-making to reduce the time, cost, and environmental footprint of R&D processes. By digitizing and automating experimental methodologies, SDLabs improve reproducibility and preserve institutional knowledge for future generations.

AI further enhances tech transfer by bridging the gap between academia and industry. It accelerates the commercialization of discoveries by optimizing experimental data, enabling scalable applications, and integrating with collaborative platforms. As a result, SDLabs and AI present a compelling vision for transforming the future of R&D and tech transfer, ensuring that innovation is both rapid and sustainable.

⁸ Schneijderberg, C. "Bureaucratization process in higher education 2017. DOI: [10.1007/978-94-017-9553-1_304-1](https://doi.org/10.1007/978-94-017-9553-1_304-1)

⁹ Kaufmann, E.; Ouschan, S. "European Academic Spin-Offs: Exploring the Barriers to Long-Term Success." 2023. [Link](#)

¹⁰ Helles, T. "Spinouts with academic founders raise less funding, study finds" 2024. [Link](#)

¹¹ Falvey, R. E., Foster, N., & Memedovic, O. "The role of intellectual property rights in technology transfer and economic growth: theory and evidence." 2016. [Link](#)

¹² Zucker, D. "Ethics and technology transfer: patients, patents, and public trust." 2011. DOI: [10.2310/JIM.0b013e318210eeb0](https://doi.org/10.2310/JIM.0b013e318210eeb0)

¹³ European Data Protection Supervisor (EDPS). "A Preliminary Opinion on data protection and scientific research" 2020. [Link](#)

2: THE ROLE OF AI AND SDLABS IN ACCELERATING R&D

Increasingly, laboratories are turning to automation and artificial intelligence (AI) to optimize the way research and development (R&D) is conducted. This new paradigm, often described as “self-driving” or “automated” laboratories (SDLabs),¹⁴ integrates a suite of digital and physical technologies—robots, sensors, high-throughput screening systems, and AI-driven analytics—to reduce experimentation time and costs while improving reproducibility (Figure 1).

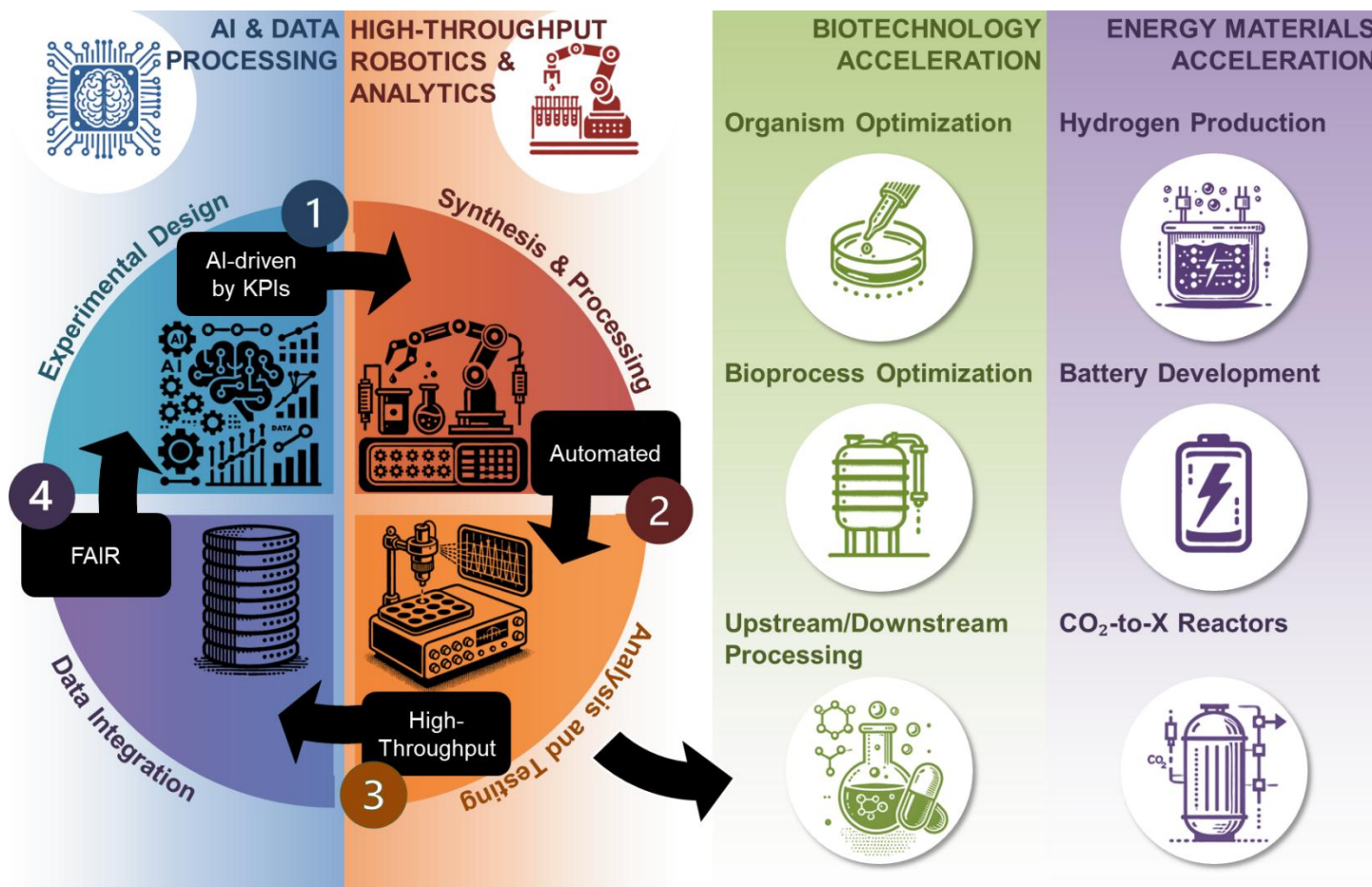


Figure 1. AI-Driven Self-Driving Laboratories for Biotechnology and Energy Materials Innovation.

2.1: DEFINING SDLABS

A Self-Driving Laboratory (SDLab) is an experimental setup that automates everything from routine sample preparation to real-time data analysis. Instead of relying solely on manual tasks, SDLabs leverage robotic platforms and sophisticated software algorithms to:

- **Design** experiments automatically
- **Execute** those experiments using robotics or other automated systems

¹⁴ “What are self-driving labs and how are they transforming the chemical industry?” Tribukait, H.; Roch, L. *World Economic Forum Annual Meeting 2024*. [Link](#)

- **Analyse** the resulting data in real time
- **Optimize** subsequent experiments based on AI-generated insights

In practical terms, researchers specify certain goals, constraints, or performance metrics. The SDLab then systematically runs a series of experiments, learns from the results, and adjusts protocols on the fly—dramatically accelerating the pace of discovery.

2.2: AI-DRIVEN DECISION-MAKING

Artificial intelligence is at the heart of SDLabs, guiding both the design and interpretation of experiments. Key AI tools include:

- **Machine Learning (ML):** Algorithms learn from past experimental outcomes to predict the most promising next steps.
- **Statistical Modelling:** Bayesian and other statistical approaches allow the lab to incorporate uncertainty and rapidly refine hypotheses as new data arrives.
- **Computational Chemistry/Biology:** Simulation tools powered by AI can predict compound properties or enzyme behaviours before real-world testing, helping narrow the range of experiments.

By linking AI models directly to automated equipment, SDLabs form a virtuous cycle of experimentation and learning: each result informs the next experiment, systematically improving outcomes in less time than traditional trial-and-error methods.

2.3: BENEFITS OVER TRADITIONAL R&D METHODS

Speed and Efficiency: Automated pipetting, high-throughput screening, and real-time data processing significantly enhance the speed and efficiency of R&D processes. These technologies minimize bottlenecks and reduce idle time, allowing for faster experimentation and data analysis.¹⁵

Cost Savings: Platforms like SDLabs help pinpoint optimal parameters more quickly and reduce the number of failed experiments. This leads to substantial cost savings and lower resource consumption. For example, in oligonucleotide synthesis, SDLabs demonstrated an 18% increase in product yield and a 22% reduction in costs.¹⁶

Reproducibility: Robotic systems can eliminate certain types of human error, and digitized processes make it easier to replicate experimental protocols and results across different labs. This enhances the reproducibility of scientific research.¹⁷

Knowledge Preservation: Digital storage of experiment designs, outcomes, and conditions helps capture institutional knowledge. This is crucial for maintaining continuity and reducing the impact of workforce transitions.¹⁸

¹⁵ Bruker “Automated Reaction and Process Development with Online Benchtop IR.” 2025. [Link](#)

¹⁶ Atinary Technologies “Maximizing yield and minimizing cost in oligonucleotide synthesis with Snapdragon Chemistry” 2024. [Link](#)

¹⁷ Holland, I.; Davies J. A. “Automation in the Life Science Research Laboratory” 2020. DOI: [10.3389/fbioe.2020.571777](https://doi.org/10.3389/fbioe.2020.571777)

¹⁸ Kamens, J. “Understanding Reproducibility: Tips for Safeguarding Experimental Outcomes in Emerging R&D Organizations” 2022. [Link](#)

Sustainability: Modern R&D practices optimize experiments to use fewer raw materials and generate less waste. They can also be designed to evaluate more environmentally friendly formulations or processes, contributing to sustainability goals.¹⁹

2.4: REAL-WORLD APPLICATIONS

Projects like the **Energy Materials Accelerator (EnerMAC)** network use the principles of SDLabs to rapidly identify and optimize new **battery**, **hydrogen**, or **CO₂-conversion** materials—enabling shorter innovation cycles for clean-tech solutions.²⁰

Similarly, the innovation network **Biotechnology Accelerator (BioAcc)**, SDLabs can test and fine-tune novel enzymes, proteins, or fermentation processes for applications in **food security** or **green manufacturing**.

The **German-Canadian Materials Acceleration Centre (GC-MAC)** leverages AI, robotics, and high-performance computing to accelerate energy materials discovery and development. By fostering collaboration between German and Canadian research institutions, it addresses key materials challenges in **electrocatalysis**, **ionic media**, **interfaces** and **electrodes**, **porous transport media**, and **materials to devices**.

2.5: FUTURE OUTLOOK

AI, robotics, and high-performance computing are converging to make SDLabs a practical reality, as evidenced by recent breakthroughs at ETH Zurich, Lawrence Berkeley National Lab, KIWI-biolab and others. Adopting SDLabs is a strategic choice that can accelerate sustainable innovation in chemistry, advanced materials, and fields pivotal for a circular economy. Leaders in industry, government, and academia hold the key to seamlessly integrating these labs for organizational and societal benefit.

Falling hardware costs and more sophisticated AI algorithms will drive SDLabs into emerging sectors like precision agriculture and additive manufacturing. Near-term gains will come from tighter connections between automated platforms and cloud data systems, boosting R&D returns. Eventually, fully autonomous, AI-driven labs—combining robotics, analytics, and modeling—could tackle challenges too complex for traditional methods, reshaping scientific inquiry.

¹⁹ Young, E. “The Importance of Sustainability in R&D” 2023. [Link](#)

²⁰ The Materials for Energy (M4E) <https://mission-innovation.net/platform/materials-for-energy-m4e/>

3: SELF-DRIVING TECH-TRANSFER FROM ACADEMIA TO START-UPS

Universities are hubs of cutting-edge research, yet many innovations remain “stuck in the lab,” never reaching the commercial sphere. Self-driving laboratories (SDLabs) and AI can serve as powerful catalysts to move scientific discoveries from academic settings into market-ready solutions. By compressing experimentation timelines, fostering data-driven workflows, and enabling new collaboration models, these tools are reshaping how emerging startups access and build upon university research.

3.1: OVERCOMING ORGANIZATIONAL BARRIERS

A major challenge in academic tech-transfer is institutional inertia. Universities typically operate with hierarchical structures and lengthy decision-making processes that can slow down the development and commercialization of new ideas. This rigidity often clashes with the rapid, iterative culture of startups.

- **Bureaucratic Hurdles:** Securing approvals and navigating legal frameworks for spin-outs can stall R&D for months or even years.
- **Risk Aversion:** Tenure and publication metrics drive faculty incentives, which may not align with entrepreneurial activities or fast-paced market launches.
- **Limited Resources:** Smaller labs may lack the funding and staffing to scale promising technologies or validate them beyond preliminary proof-of-concept work.

3.2: FOSTERING AN ENTREPRENEURIAL CULTURE

Transforming academic institutions into innovation engines requires a cultural shift toward entrepreneurship and agility.

- **Accelerator Programs:** University-based incubators or accelerator programs can streamline the path from initial idea to startup formation. These programs encourage rapid prototyping and early-stage fundraising—essential for tech-intensive ventures that need to demonstrate quick results.
- **Agile Research Entities:** By establishing smaller, flexible research units that function like startups, universities can reduce bureaucratic overhead and focus on real-world applications of research.
- **Collaborative Platforms:** Partnerships with large development agencies or cross-institutional programs—such as the [**WFP Innovation Accelerator**](#) model—showcase how startups can tap into big data resources, domain expertise, and broader networks without losing their innovative edge.

*The companies that will create the most economic value in the future will be the ones that find ways to **participate more effectively in a broader range of more diverse knowledge flows** that can refresh knowledge stocks at an accelerating rate.*

John Hagel
Deloitte Center for the Edge

3.3: THE ROLE OF SDLABS AND AI IN TECH-TRANSFER

Self-driving labs and AI-driven experimentation help bridge the gap between academic research and industry needs:

1. **Speed to Market:** Automated experimentation compresses months of manual bench work into days or weeks, giving startups a competitive edge in time-sensitive fields like drug discovery or advanced materials.²¹
2. **Scalability:** SDLabs can run large volumes of tests simultaneously, allowing new ventures to stress-test various formulations or prototypes at minimal cost.
3. **Data-Driven Insights:** AI-based models can reveal hidden patterns or predict successful outcomes, enabling startups to fine-tune products before making large-scale investments.
4. **Lower Risk:** Rigorous, reproducible experiments reduce uncertainty for investors, partners, and potential customers, facilitating smoother fundraising and partnerships.

3.4: CASE STUDIES

CASE 1: NxtGen Science (NGS) [DE] Accelerated Innovation and Data Valorization

NxtGen Science (NGS) is pioneering a platform where academic research data generated via SDLabs can be curated and made available to industry partners. By applying **FAIR principles** (Findable, Accessible, Interoperable, Reusable) and leveraging **IP-NFTs** (intellectual property non-fungible tokens) for secure data rights, NGS creates a marketplace where academia and industry can:

- **Acquire** high-quality, reproducible experimental datasets without incurring high in-house R&D costs.
- **Commission** customized, crowd-funded research tailored to their unique product goals.
- **Validate** hypotheses rapidly, thereby accelerating the path from lab prototype to commercial viability.

IP-NFTs are digital tokens representing intellectual property, created by the host-platform. They combine blockchain technology with legal contracts, enabling secure and transparent ownership, funding, and trading of scientific discoveries. IP-NFTs simplify the commercialization of research by decentralizing IP rights and linking them to smart contracts.
<https://docs.molecule.to/documentation/ip-nfts/intro-to-ip-nft>

CASE 2: ChimiaDAO [CO] Decentralizing Tech-Transfer

ChimiaDAO integrates blockchain technology to establish transparent, tamper-proof records of scientific findings and transactions. Through a decentralized network, it:

²¹ "How AI is transforming the factory floor" Schönfuß, B. *World Economic Forum Annual Meeting 2024*.
<https://www.weforum.org/agenda/2024/10/ai-transforming-factory-floor-artificial-intelligence/>

- **Verifies Experimental Data:** Immutable blockchain ledgers ensure that academic discoveries remain trustworthy and unaltered, giving startups confidence when licensing or investing in emerging technologies.
- **Incentivizes Sharing:** Researchers are rewarded with tokens for contributing valuable data, fostering collaboration across institutions and disciplines.
- **Reduces Friction:** By automating aspects of intellectual property management, ChimiaDAO accelerates negotiations and licensing deals—crucial steps in bringing novel science to market.

3.5: PATH FORWARD

By leveraging SDLabs, AI, and decentralized collaboration platforms, universities and startups can rapidly co-develop breakthroughs that address industrial and societal challenges. To seize these opportunities, academic leaders should **adopt lean practices** in university labs to shorten development cycles and quickly validate market needs. They can **strengthen IP frameworks** that fairly reward both researchers and industry partners, encouraging shared investment in innovation. **Cross-sector collaborations** with venture capitalists and established companies will further amplify resources, expertise, and funding. By aligning incentives, modernizing research structures, and integrating automation and AI, academia can transform into a powerful launchpad for high-impact ventures—driving economic growth, generating jobs, and advancing meaningful technologies.

4: CONTRACT RESEARCH 4.0

Outsourcing has become a pivotal strategy for companies aiming to reduce overhead, accelerate product development, and leverage specialized expertise. In both established and emerging industries—from pharmaceuticals and biotechnology to automotive engineering—Contract Research Organizations (CROs) play a critical role in bridging innovation gaps, providing flexible R&D capacity, and offering niche skill sets. As AI, big data, and self-driving laboratory (SDLab) technologies evolve, the next phase of contract research—“Contract Research 4.0”—promises not only faster workflows but also entirely new business models that prioritize collaboration and data valorization.

4.1: EVOLVING CRO DYNAMICS

Increasing R&D Costs and Complex Projects. R&D in fields like advanced therapeutics or alternative energy often demands significant capital and specialized infrastructure. CROs can offer turnkey solutions, from early-stage discovery to clinical testing, helping companies bypass the costs and risks of building these capabilities in-house.

Speed-to-Market Pressures. Tightening patent windows and fierce global competition drive companies to shorten development cycles. CROs that integrate automated experimentation and AI-enabled analytics can dramatically cut trial times, positioning themselves as indispensable partners in race-to-market environments—particularly in life sciences and green tech.

4.2: REGIONAL SPOTLIGHTS

Germany

- **High Investment in External R&D:** Germany allocated ~\$29.5 billion to external R&D in 2022, marking a 4.1% increase from the previous year. Around 35 cents of every euro spent on internal R&D goes to outsourcing.
- **Sectoral Leaders:** Pharmaceuticals, automotive, IT, and machine-building are among the top consumers of CRO services. Many mid-sized enterprises (Mittelstand) rely on contract research for specialized tasks they cannot handle in-house, especially in areas like gene therapy or sustainable materials.
- **Growth Drivers:** Cost efficiency and access to niche expertise remain central. AI and robotics are increasingly integrated into lab workflows, transforming the speed and scale of experimental research.

United Kingdom

- **Global Player in Life Sciences:** The UK invested ~\$5.5 billion in CRO services in 2022, with a strong focus on pharmaceuticals, biotech, and clinical trials. Key demand drivers include oncology and metabolic research (e.g., obesity treatments).
- **Patent Cliffs and AI Adoption:** Impending patent expirations around 2029 spur companies to innovate quickly. AI-driven CROs offer accelerated trial designs, predictive modeling, and more robust post-market surveillance.

- **SME Involvement:** Smaller biotech startups look to CROs to manage specialized tasks like gene editing and next-gen sequencing—domains that require both advanced labs and cutting-edge data analytics.

Table 1. Contract Research in Germany and the UK.

	Germany	UK
Contract Research Investments	\$29.53 billion in 2022, with 4.1% growth.	\$5.48 billion in 2022 for CRO services.
Sector Focus	Pharma, automotive, IT, machine-building sectors.	Pharma, biopharma, oncology, and obesity treatments.
SME Involvement	High involvement from the Mittelstand and SMEs.	High reliance on CROs by emerging biopharma firms.
AI/ML Adoption	Increasing use in pharma and automotive sectors.	Widespread adoption in life sciences and clinical trials.
CRO Market Size	Part of the \$29.53 billion external R&D investment.	\$19.68 billion in 2022, planned to grow \$39.55 billion by 2030 (9.0% CAGR).
Growth Drivers	Cost-efficiency, specialized expertise in R&D.	Patent cliff in 2029, fast-tracked drug development, AI integration.

4.3: "THE INNOVATION AMPLIFICATION LOOP"

Currently, contract research accounts for approximately 33% of the revenue generated by some of the Berlin University Alliance's (BUA) tech-transfer offices (TTOs). While this revenue stream already supports research and knowledge transfer, the Innovation Amplification Loop demonstrates how strategic shifts—particularly Contract Research 4.0 and IP-NFTs—can substantially improve funding for publicly funded universities and private contract research organizations (CROs) (Figure 2).

In this loop, universities form deeper partnerships with industry and startups, moving beyond one-off projects to collaborate on multi-year data-sharing and commercialization strategies. This approach not only bolsters the existing revenue baseline but also ensures a steady influx of funds for cutting-edge R&D. **Specialized or decentralized SDLabs** then rapidly transform funded ideas into tangible results, aided by **FAIR data standards** that make outputs widely accessible and highly reusable. **Blockchain-based IP-NFTs** further accelerate this process by simplifying licensing and providing transparent ownership records, enabling faster technology transfers and clearer royalty agreements.

As startups emerge from these initiatives, they leverage ready-to-use data and secure IP frameworks to bring products to market quickly. Their commercial success in turn feeds back into the academic ecosystem, creating an upward spiral of innovation, investment, and societal impact. Over time, these combined elements significantly enhance the financial sustainability of both universities and private CROs, transforming contract research from a useful revenue source into a strategic engine driving large-scale research and development.

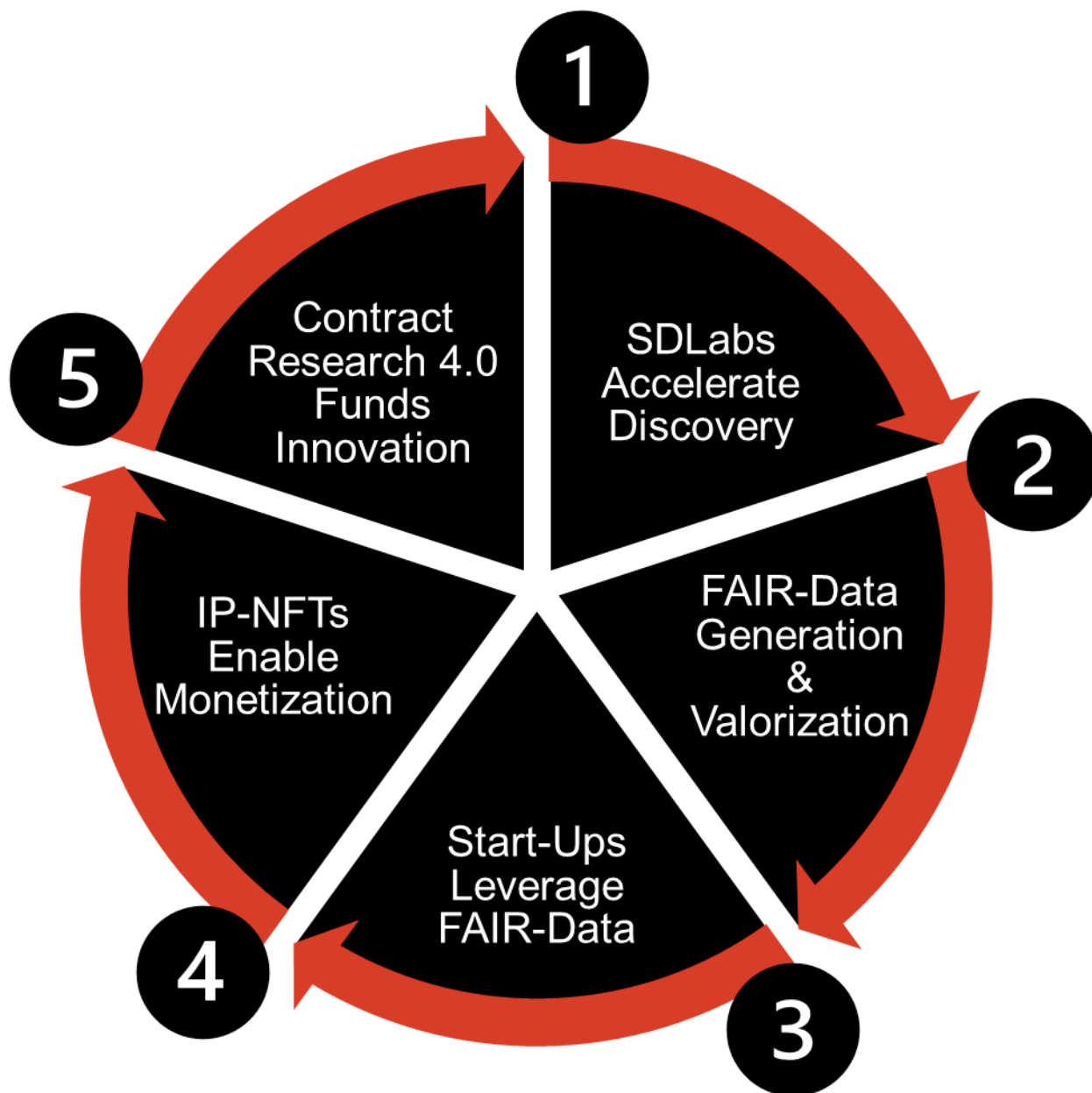


Figure 2. The Innovation Amplification Loop: A self-reinforcing innovation cycle that flows from contract research funding through accelerated discovery, FAIR data generation, startup growth, and IP-NFT-based monetization.

5: OPEN & FAIR SCIENCE: ENABLING COLLABORATION AND INNOVATION

As scientific and technological challenges grow more complex, traditional, siloed R&D models become less effective. Enter the paradigm of Open & FAIR Science—an approach defined by **transparency**, **collaboration**, and **equitable access** to research outputs. By adhering to the FAIR principles (Findable, Accessible, Interoperable, and Reusable), and embracing open collaboration frameworks, researchers worldwide can more rapidly generate, validate, and refine scientific knowledge.

5.1: THE CASE FOR OPEN SCIENCE

Transparency and Trust. Open-science practices like open-source code, publicly accessible data repositories, and pre-prints cultivate trust in scientific findings. This openness helps mitigate reproducibility concerns and fosters peer verification.

Accelerated Innovation. When researchers—across academia, industry, and governmental organizations—share data and insights in near real-time, discoveries can be rapidly validated and adapted for real-world applications. This is especially powerful in fields like drug development and sustainability research, where global challenges demand swift, collective action.

Inclusive and Equitable Collaboration. Open repositories and FAIR-aligned data reduce geographical and institutional barriers. Scientists from under-resourced regions or smaller startups can access the same foundational data as large corporations or well-funded labs, thereby democratizing innovation.

5.2: CURRENT TRENDS IN OPEN & FAIR SCIENCE

The landscape of Open and FAIR Science is rapidly evolving, driven by technological advancements, global policy initiatives, and increasing societal demands for transparency and accessibility. Several key trends highlight the progress and challenges in this field:

1. Integration of Artificial Intelligence (AI) in Scientific Research AI is a catalyst for innovation in Open Science, offering tools to accelerate data analysis, automate repetitive tasks, and generate new hypotheses. In drug discovery, for instance, AI has significantly reduced the time required for identifying potential candidates, with applications such as predictive toxicology enhancing safety protocols and environmental sustainability. AI-powered systems like ChestX-Transcribe demonstrate the potential for automated medical report generation, improving both efficiency and accuracy.

FAIR is an acronym for **Findable, Accessible, Interoperable, and Reusable** to guide data-sharing strategies:

Findable: Assign unique IDs and relevant metadata for easy discovery.

Accessible: Use clear protocols, via open or controlled repositories, to make data readily obtainable.

Interoperable: Adopt standardized formats and metadata for seamless cross-platform integration.

Reusable: Provide detailed documentation and licenses to facilitate replication and extension.

However, ensuring the quality and fairness of AI outcomes necessitates robust FAIR-compliant datasets, as biases and inconsistencies can undermine scientific integrity.²²

2. Expansion and Application of FAIR Principles The FAIR Principles, introduced in 2016, continue to evolve, emphasizing not only data findability but also true discoverability and accessibility. Challenges remain in ensuring consistent metadata descriptions and overcoming barriers posed by non-open repositories. As highlighted in recent policy discussions, extending FAIR principles to enable AI-ready and cross-domain interoperable datasets will be pivotal for scaling Open Science globally.²³ Enhanced standards for interoperability and data integration across disciplines are increasingly being adopted as part of international research collaborations.

3. Development of Sustainable Research Infrastructures Infrastructure development is critical to the success of Open Science. The ESFRI Roadmap 2026 illustrates the need for financial and operational sustainability in research infrastructures. By incorporating environmental considerations and long-term planning, the roadmap sets a precedent for integrating sustainability into the research ecosystem.²⁴ Furthermore, global partnerships and the coordination of distributed research infrastructures ensure that Open Science initiatives remain resilient and effective in addressing diverse scientific challenges.

4. Ethical and Regulatory Advances in Data Governance Data governance frameworks are evolving to address ethical concerns, including privacy, security, and equitable access. The need for clear regulations has been underscored by challenges such as the misuse of data and the ethical implications of AI-driven decisions. Initiatives like the European Open Science Cloud (EOSC) emphasize the importance of creating a unified and transparent data-sharing environment that aligns with global standards.²⁵

5. Increasing Global Collaboration The global nature of scientific challenges requires collaborative approaches that transcend national and disciplinary boundaries. Research infrastructure networks and data-sharing platforms are fostering international partnerships, enabling the pooling of resources and expertise. The focus on global interoperability standards and mutual agreements has positioned Open and FAIR Science as a cornerstone of international scientific cooperation.²⁶

5.3: THE FUTURE OF COLLABORATION

Open and FAIR Science is rapidly advancing alongside emerging technologies such as self-driving labs, digital twins, and federated learning. Together, these innovations enable real-time research networks that reduce duplication and accelerate discoveries across continents. Continuous experimentation guided by cloud-based AI and automated lab platforms will further integrate global data updates and shifting research priorities. At the same time, new business models—like data marketplaces and decentralized autonomous organizations (DAOs)—are poised to reshape how collaborative research is funded and rewarded. By embracing open practices, consistent data standards, and ethical guidelines, stakeholders can unlock reproducible breakthroughs that drive transformative, transparent progress.

To maximize the potential of this paradigm, **international collaboration should be strengthened**. Cross-border partnerships enhance resource sharing and knowledge exchange, as demonstrated by

²² Wilkinson, M. D., et al. "The FAIR Guiding Principles for scientific data management and stewardship." *Scientific Data*, 3(1), 2016. DOI: [10.1038/sdata.2016.18](https://doi.org/10.1038/sdata.2016.18)

²³ European Commission. "European Open Science Cloud (EOSC): Strategic Priorities." 2023. [Link](#)

²⁴ ESFRI. "ESFRI Roadmap 2026: Strategy and Vision for European Research Infrastructures." 2024. [Link](#)

²⁵ World Economic Forum. "AI for Scientific Discovery: Transforming Research Methodologies." 2025. [Link](#)

²⁶ Global Science Forum (OECD). "International collaboration in science." 2024. [Link](#)

initiatives like the European Open Science Cloud (EOSC) and global research infrastructure networks.²⁷ Governments, industry, and academic institutions must also **expand investments in robust, FAIR-compliant research infrastructure**—spanning data repositories to AI-ready systems—to ensure equitable access and interoperability. **Consistent metadata standards** are equally crucial; by harmonizing approaches across scientific domains, researchers can improve data discoverability and integration.²⁸ Moreover, **cultural change within the scientific community** is vital: academic institutions and funding agencies should revise evaluation metrics to recognize and reward contributions to data sharing and collaborative research.

AI can play a pivotal role in accelerating FAIR Science by automating data integration, boosting metadata quality, and identifying gaps in existing research. By **leveraging these tools and addressing regional disparities**—through training programs, infrastructure development, and capacity-building—stakeholders can bridge the digital divide and incorporate diverse perspectives into global research efforts.

Ultimately, by pursuing these strategies and investing in Open and FAIR Science, the global research ecosystem can become more inclusive, equitable, and innovative, ushering in a future defined by transparency, reproducibility, and transformative impact.

5.4: FUTURE OUTLOOK: EMERGING MODELS AND INFRASTRUCTURE

The convergence of artificial intelligence, automation, and collaborative research networks is shaping a bold new era for scientific discovery and technology development. Governments and industry stakeholders worldwide are making unprecedented investments in high-performance computing (HPC) facilities, decentralized data platforms, and large-scale AI projects. These new infrastructures are set to redefine not only how research is conducted, but also who participates, and how results are translated into societal impact.

CASE 1: US Stargate Project

In the United States, the private-sector-led Stargate Project is deploying billions of dollars to build a coast-to-coast network of AI-centric data centers. By collaborating with major technology firms, it seeks to provide vast computing capacity for cutting-edge applications—from genomics to climate modeling. While such monumental infrastructure promises transformative capabilities, questions persist around equitable access, transparency, and oversight in the allocation and use of these resources.²⁹

CASE 2: EU AI Factories

By contrast, the European Union is launching publicly co-funded “AI Factories” that leverage its network of HPC supercomputers. Structured around principles of responsible AI, these initiatives focus on critical sectors like healthcare, clean energy, and advanced manufacturing. Through open standards, shared data repositories, and regulatory guardrails, the EU model aims to foster cross-border collaboration and ethical AI practices.³⁰

²⁷ EOSC Association. "Cross-domain recommendations and feedback for the EOSC Interoperability Framework." TBA (Q2 2025). [Link](#)

²⁸ FAIRsFAIR. "Recommendations on practice to support FAIR data principles" 2020. DOI: [10.5281/zenodo.3924131](https://doi.org/10.5281/zenodo.3924131)

²⁹ NBC News. "Trump announces 'Stargate' AI infrastructure project with 'colossal data centers'" 2025. [Link](#)

³⁰ European Commission. "AI Factories" 2024. [Link](#)

Aspect	US Stargate Project	EU AI Factories
Objective	Build a \$500 billion AI data center network for global leadership in AI, with a focus on infrastructure and innovation.	Establish AI factories leveraging European HPC supercomputers to ensure ethical and collaborative AI development.
Scale of Investment	\$500 billion over four years (private-led, funded by OpenAI, SoftBank, Oracle, MGX).	€2 billion (€1 billion EU-funded, €1 billion from Member States).
Geographic Scope	Nationwide buildout starting in Texas, with a focus on the US market.	Distributed across the EU, forming a networked, collaborative ecosystem.
Tech Partners	Microsoft, NVIDIA, Oracle, Arm, OpenAI.	Supported by EuroHPC supercomputers, Digital Innovation Hubs, and Testing and Experimentation Facilities.
Focus Areas	AI infrastructure for OpenAI, focusing on computing capacity, innovation, and geopolitical leadership.	AI applications in strategic sectors: healthcare, energy, transport, defence, and manufacturing.
Ethics and Trustworthiness	Minimal public focus on ethical guidelines; driven by economic and strategic goals.	Strong emphasis on ethical AI via the EU AI Act and safety regulations.
Implementation Timeline	Buildout underway; initial completion within four years.	Rolling project setup until 2025; first deadline November 2024.
Strategic Goals	US dominance in AI, economic leadership, and technological innovation.	European collaboration, ethical AI, and leadership in high-impact sectors.

Author

Prof. Dr. Michael J. Bojdys

Consultant and Professor in Materials Chemistry



WEB www.bojdyslab.org

TEL +491774818190

EMAIL m.j.bojdys.02@cantab.net

bojdysLAB

Center for the Science of Materials Berlin
(CSMB) der Humboldt-Universität zu Berlin
Zum Großen Windkanal 2
12489 Berlin
Germany

Disclaimer

This document, "*AI and SDLabs for Scientific Discovery & Tech-Transfer*", was prepared by Michael J. Bojdys and bojdysLAB to offer insights into the role of AI and self-driving laboratories (SDLabs) in advancing scientific discovery and technology transfer. The content reflects a collaborative effort but may not represent the views of all contributors. While every effort has been made to ensure the accuracy and completeness of the information, no guarantees are provided. Portions of this document have been machine-generated and/or machine-translated.

© 2024 Michael J. Bojdys and bojdysLAB. All rights reserved. No part of this publication may be reproduced or transmitted in any form or by any means, including photocopying and recording, or by any information storage and retrieval system.