

FOXDEN: Advancing AI-Ready Data Exchange Through CHES–MagLab Collaboration

[David S. Butcher](#), [Lissa C. Anderson](#)

[Ion Cyclotron Resonance Facility](#), [National High Magnetic Field Laboratory](#), Tallahassee FL

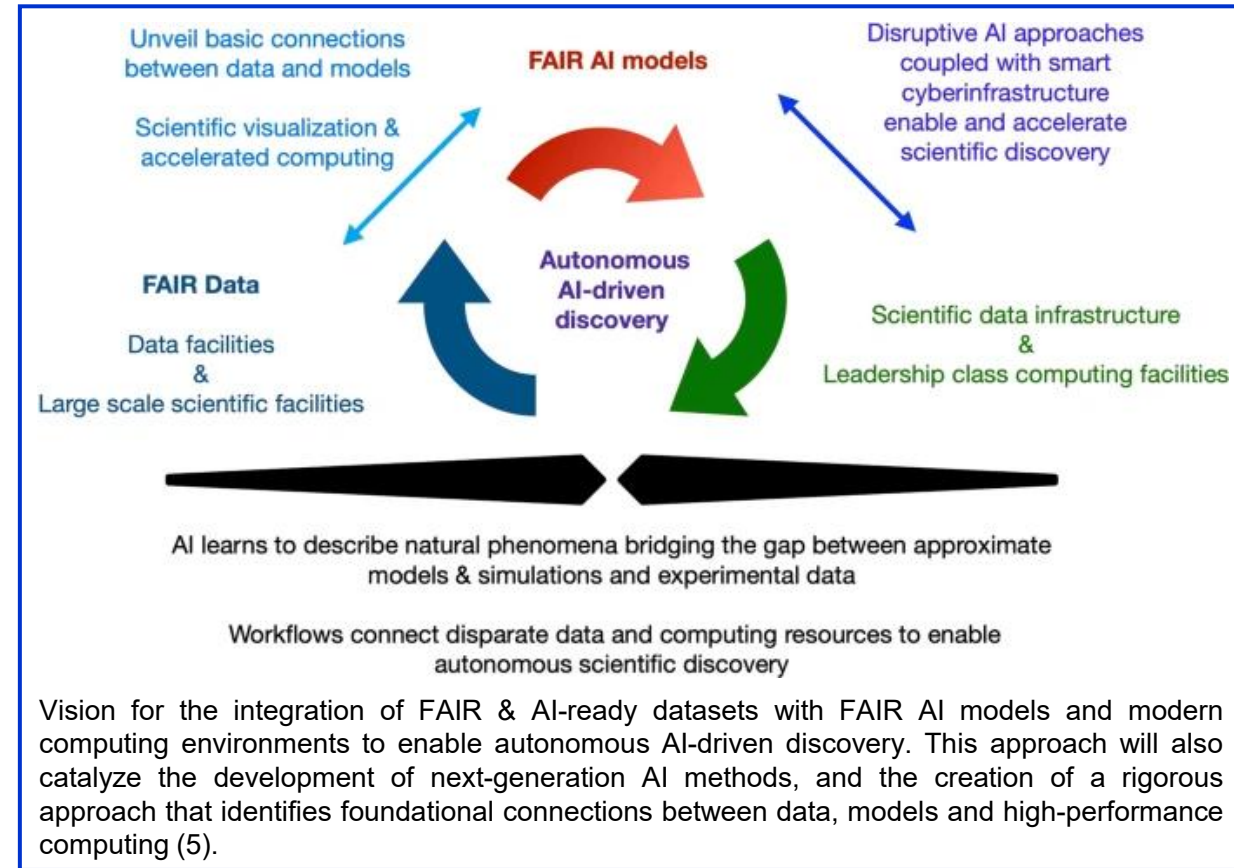


Funding Grants: K. M. Amm (NSF DMR-2128556)

Over the past decade, the concept of FAIR data—**Findable, Accessible, Interoperable, and Reusable** for both humans and machines—has become a cornerstone of modern scientific research. FAIR principles are essential for ensuring reproducibility, integrity, and rigor across disciplines. With the rapid rise of artificial intelligence (AI) tools in research, the **machine-actionable** aspect of FAIR data has gained new urgency: data that is FAIR is also **AI-ready**, enabling responsible deployment of AI agents to accelerate discovery. AI-ready data is foundational to emerging innovations such as **self-driving laboratories**, which can conduct fully automated experiments with minimal human intervention (1).

Despite its importance, achieving FAIR data remains a challenge in fields where detailed metadata collection is not standard practice. To address this, the **National High Magnetic Field Laboratory (MagLab)** is collaborating with researchers at the **Cornell High Energy Synchrotron Source (CHESS)** to implement new cyberinfrastructure solutions. CHESS has developed the **FAIR Open-Science Extensible Data Exchange Network (FOXDEN)** (2) — a suite of services for managing data and metadata, visualizing results, and publishing to external repositories.

The MagLab is building custom applications on FOXDEN to automate the generation of rich metadata using emerging standards such as **Science-on-schema.org** (3) and **Croissant** (4). These tools enhance dataset discoverability and support AI and machine learning workflows. By integrating these innovations, the MagLab is making FAIR and AI-ready data more accessible to its user community — empowering researchers to harness advanced computational methods and contribute to a more connected, intelligent research ecosystem.



Citations: (1) Canty, R. B. *et al.* Science Acceleration and Accessibility with Self-Driving Labs. Nat Commun 2025, 16 (1), 3856. <https://doi.org/10.1038/s41467-025-59231-1>.

(2) CHESS Computing: FOXDEN. <https://github.com/CHESSComputing/FOXDEN>

(3) Science-on-Schema.Org v1.3.2, 2024. <https://doi.org/10.5281/ZENODO.7884538>.

(4) Akhtar, M. *et al.* Croissant: A Metadata Format for ML-Ready Datasets. <https://doi.org/10.1145/3650203.3663326>.

(5) Ravi, N. *et al.* FAIR principles for AI models with a practical application for accelerated high energy diffraction microscopy. Sci Data 2022, 9, 657. <https://doi.org/10.1038/s41597-022-01712-9>