

## Status and perspectives of ILDG

---

**Christian Schmidt**<sup>a,\*</sup>

<sup>a</sup>*Fakultät für Physik, Universität Bielefeld,  
Universitätsstrasse 25, 33615 Bielefeld, Germany*

*E-mail:* [schmidt@physik.uni-bielefeld.de](mailto:schmidt@physik.uni-bielefeld.de)

We discuss the status and progress of recent efforts to modernize the International Lattice Data Grid (ILDG). This includes activities of the metadata and middleware working groups concerning deployment and operation of crucial services (user management, metadata catalogues, file catalogues) and extensions of the metadata format, which have been tailored according to the needs of the large collaborations. We also report on developments and extensions that are planned to be addressed in the foreseeable future.

*The 42nd International Symposium on Lattice Field Theory (LATTICE2025)  
2-8 November 2025  
Tata Institute of Fundamental Research, Mumbai, India*

---

\*Speaker

## 1. Introduction

The International Lattice Data Grid (ILDG) [1–5] is an initiative with the purpose of enabling and coordinating the sharing of research data within the research community in Lattice Field Theories. The ILDG is organized as a Virtual Organization (VO). It provides a central user management service [6] and defines the specifications of the metadata schemata, the file format, and the interface of the catalog services. In turn, the actual storage resources and catalog services are provided and managed in an autonomous way by the various regional grids. Traditionally, the grids have been CSSM [7] for Australia, JLDG [8] for Japan, Latfor Data Grid (LDG) [9] for continental Europe, UK Lattice Field Theory for the UK [10], and USQCD [11] for the US. The ILDG has two working groups, the metadata working group (MDWG), which specifies and develops the metadata standard and file format, and the middleware working group (MWWG), which defines and coordinates the implementation of services and middleware for the interoperability of the regional grids. These activities are overseen by the ILDG board. Since 2021, the ILDG benefits from the German research data infrastructure initiative (NFDI). Within one of its consortia the Particles, Universe, NuClei and Hadrons for the NFDI (PUNCH4NFDI), the java code for the catalog services has been refactored.

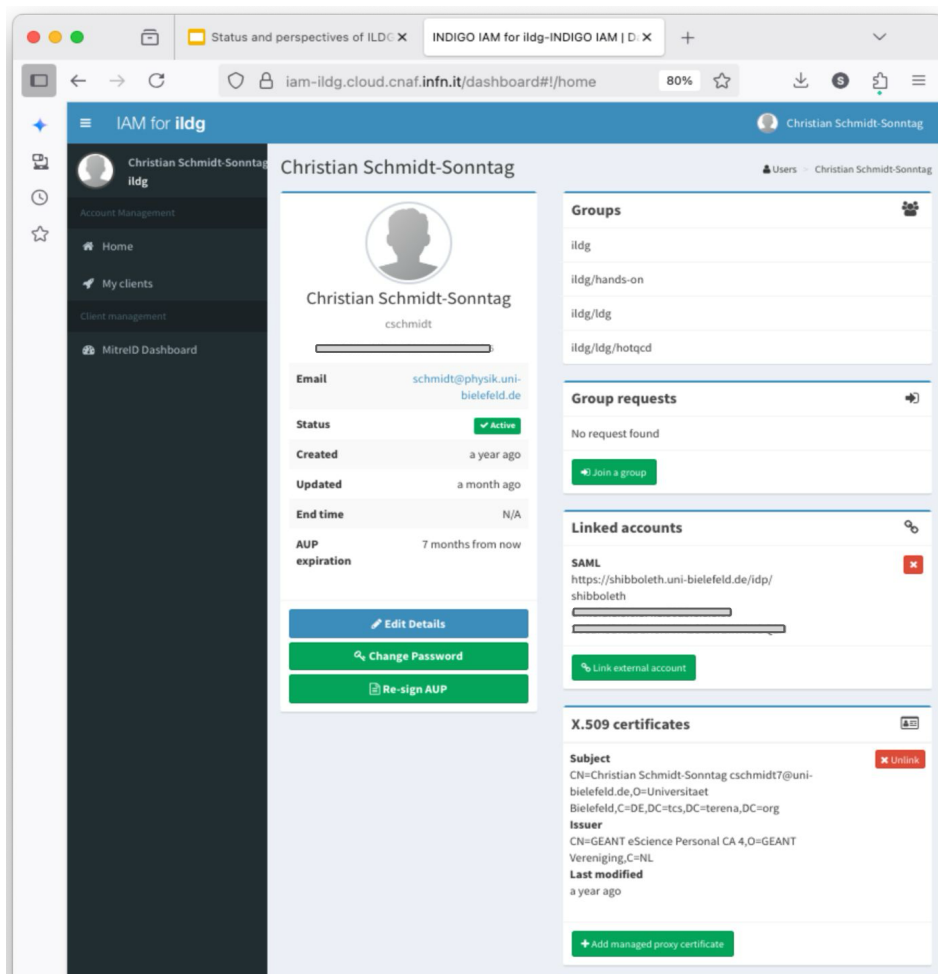
## 2. Recent highlights

The International Lattice Data Grid (ILDG) has recently completed a major upgrade and modernization, referred to as ILDG 2.0, which is now fully operational and ready for use. Key components, like the federated user management service and the metadata catalogs have been replaced or modernized, and the metadata schema has been extended. The important changes are discussed below.

### 2.1 The Identity and Access Management (IAM)

Users and VO membership are now managed through the INDIGO Identity and Access Management (IAM) service, which is developed and maintained by INFN-CNAF. The dedicated IAM instance for ILDG is also hosted at CNAF and currently there are  $O(100)$  ILDG members registered in the IAM. It is configured with Single-Sign-On, using the eduGAIN federation of identity providers (IdP). This means that all users at login to the IAM are redirected to the IdP of their home institutions, where they can authenticate with their local credentials. This completely eliminates the need for users to hold a valid Grid certificate, which in some cases was a major obstacle for using the ILDG. Also authentication and authorization towards the services is now fully based on tokens. With this significantly simplified user registration and a fine-grained token-based access control, two major milestones in the ILDG agenda have been reached.

When requesting ILDG membership at the IAM, users are required to accept the VO policy [12] and an Acceptable Use Policy (AUP) [13]. To retain the membership, the acceptance has to be repeated periodically every twelve months. After login to the IAM, the user will see a dashboard, as shown in Fig. 1. Besides information on the user's identity and the button to re-sign the AUP, the user is able to create clients that can receive capability tokens from the IAM. In turn, these tokens provide authorization to access storage elements for reading or writing of configurations, or

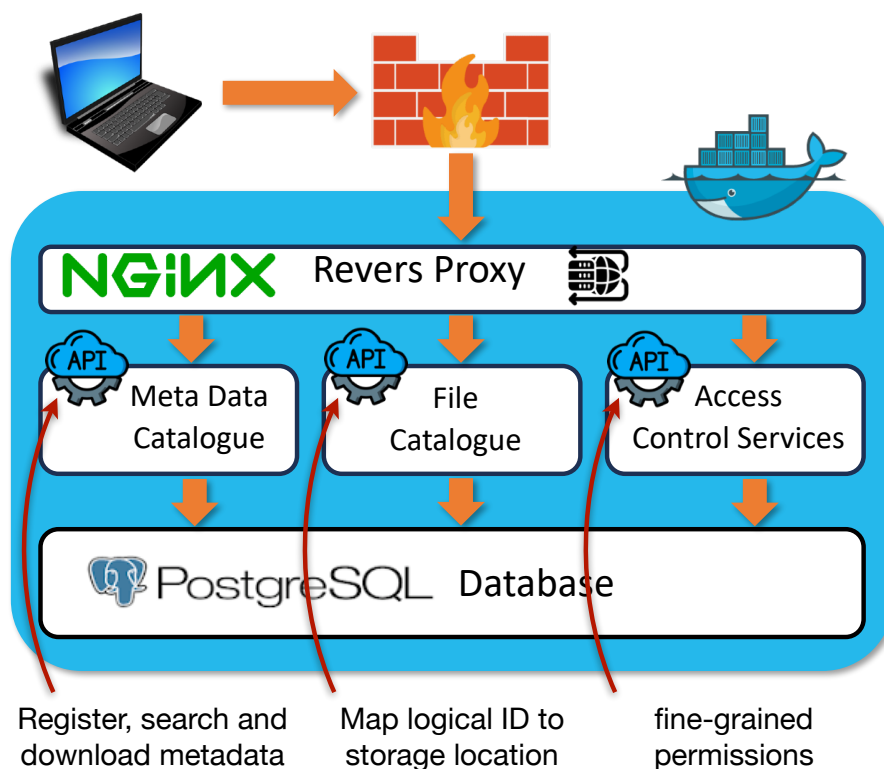


**Figure 1:** Dashboard of an ILDG user registered in the IAM

to register metadata in the catalogs. The scope of the tokens defines the detailed read- and write permissions of the user for specific resources, e.g. in certain directories. The allowed scopes can be derived from the group memberships of the user, which are also managed in the IAM and visible in the dashboard. Finally, we note the possibility to link the account of the user to an X.509 grid certificate, which can then be used as an alternative and convenient authentication method for the IAM login.

## 2.2 Catalog services

ILDG uses two kinds of catalogs: metadata catalogs contain all metadata that describes the ensembles and configurations, while file catalogs map the logical ID of configurations to their (possibly replicated) storage locations. The interface specifications of the metadata and file catalogs in ILDG 2.0 have been updated to a modern (REST) API. They include complex search operations and downloading of metadata, as well as optional extensions for write-access with fine-grained token-based authorization. A corresponding reference implementation of the catalogs, which can



**Figure 2:** Structure of a containerized implementation of ILDG web services.

also be deployed as stand-alone services, is available as Docker containers. Their general structure is shown in Fig. 2. Containerization and making use of the same implementation of the catalogs, as currently done by the LDG, JLDG and UKQCD regional grids, considerably simplified the deployment and operation of these services, and it may help to join efforts for their long-term maintenance.

An additional access control service can be used to define very fine-grained authorization policies in a hierarchical way. Currently, this feature is not yet used and authorization decisions are directly managed by scope policies defined within the IAM. The coherent and fully token-based access control to all (meta-)data resources is an important new feature in ILDG 2.0, because it enables collaborations to keep uploaded data under embargo for some limited amount of time, i.e. to reliably restrict access to the configurations to internal use only, before they are made public.

We also note that the new implementation of the metadata catalog is completely agnostic to the used metadata schemata. This makes it an attractive solution also for other communities. As part of the work program in the second funding period of PUNCH4NFDI it is planned to realize ILDG-like data management frameworks also for other scientific communities.

### 2.3 Metadata schema and file format

A revised and extended version 2.0 of the QCDml metadata schemata for ensembles and configurations has been released. It incorporates the requirements from several lattice QCD collaborations and enables them to markup their configurations according to their needs. Most importantly, the

schema is now fully FAIR compliant [14], where FAIR stands for "Findable, Accessible, Interoperable and Reusable". It now includes a mandatory license statement, optional funding information, and the possibility to indicate embargo periods. Several new lattice actions, e.g. QCD+QED, gauge groups beyond SU(3), and open boundary conditions are now also supported. In addition, the format specifications for data files have been extended to support a compressed data format and packing of multiple configurations into a single LIME file.

### 3. Interacting with ILDG services

Simple command line tools to interact with the catalogs for the purpose of uploading, searching, and downloading of metadata are available as repository or within a container [15]. They are basically wrapper scripts using *curl* [16] commands to generate suitable HTTP requests and do not require installation of any complex grid software. More high-level tools to perform the entire upload workflow with all required write operations to metadata catalog, storage elements, and file catalog are still desirable (possibly also for download operations, although these are simpler). With the API of the catalogs being documented (and accessible) through the Swagger interface [17], it is also straight forward for users to develop their own customized tools, web pages, or scripts for the interaction with ILDG services. Several examples of web interfaces are now online, ranging from simple listings [18] to the reactivated faceted navigation page of JLDG and the newly developed markup and search interfaces [19, 20]. An add-on service for metadata harvesting via OAI-PMH (Open Archives Initiative Protocol for Metadata Harvesting) is also available to integrate ILDG data into other research data services, like INSPIRE or the PUNCH science data platform.

Storage elements are accessed through standard protocols (typically also HTTP or WebDAV) with a broad choice of existing client tools. ILDG currently has a total of six storage elements in the various regional grids: four in LDG (DESY Hamburg and Zeuthen, Jülich Supercomputing Centre, and INFN-CNAF), one in Japan (Tsukuba University) and one in the UK (Swansea University). They are configured to support at least read access with tokens issued by the IAM of ILDG. In total about 400 ensembles with 350,000 configurations (mostly legacy data from ILDG 1.0) are currently stored on the ILDG storage elements. Various collaborations have plans for massive uploads of new ensembles to ILDG, see Ref. [21, 22].

### 4. Outlook

ILDG 2.0 with its revised and modernized services, infrastructure, and metadata schema is now ready to be used for new data. Various collaborations have plans for massive uploads of new ensembles to ILDG, see Ref. [21, 22]. Starting such uploads requires some time and preparatory work by the data providers. However, this effort can be expected to quickly pay off through an improved (even only collaboration internally) and fully FAIR-compliant data management for the gauge configurations.

Further development efforts by the working groups and user community are expected to focus on more convenient and high-level client tools with graphical as well as command-line interfaces. Moreover, a tight monitoring of the status of all services (IAM, catalogs and storage) is desirable for a highly reliable production mode. Development plans of the metadata working group include the

support for additional file formats, such as HDF5, and the setup of standardized workflows for data publishing, including DOI minting, landing pages, and metadata harvesting. Community training and on-boarding activities, like the recent hands-on workshop [23], need to be repeated for both data providers and as data consumers.

Looking forward, joint efforts are needed to ensure long-term sustainability of ILDG activities through adequate funding and manpower commitments, and the lattice community might want to reassess the structure of regional grids. In Europe, sustainability questions are currently further addressed and organized through EuroLFT [24]. In addition, ILDG aims to broaden the impact of its technology beyond lattice QCD towards other scientific communities. These activities are pursued, for instance, within the PUNCH4NFDI consortium and might generate synergies on efficient and FAIR compliant research data management. Continued participation in ILDG working groups is encouraged to plan and realize future developments. For questions or suggestions please contact the ILDG working groups or board through `ildg-contact(at)desy.de`.

## Acknowledgments

The status and progress of ILDG 2.0 as reported here has only been possible through the great effort of members of the ILDG working groups and the support by the ILDG board. In particular, we point out recent essential contributions and work by Antonio Rago, Basavaraja Bheemalingappa Sagar, Craig McNeile, Daniel Knüttel, Dirk Pleiter, Ed Bennet, Georg von Hippel, Giannis Koutsou, Giovanni Pederiva, Hideo Matsufuru, Hiroshi Ohno, Hubert Simma, Osamu Tatebe, and Yoshinobu Kuramashi. A crucial element of ILDG 2.0 is the INDIGO IAM, and we are grateful to the INFN-CNAF team for their great development, support, and help, in particular to Federica Agostini, Francesco Giacomini, Roberta Miccoli, Diego Michelotto, Carmelo Pellegrino, and Enrico Vianello. We thank the IT teams of all ILDG sites for their support and help, in particular, at CNAF, DESY, Jülich Supercomputing Centre, and Tsukuba University. We acknowledge important support from various members of the PUNCH4NFDI consortium, in particular, Gunnar Bali, Olaf Kaczmarek, Frithjof Karsch, Stefan Krieg, Carsten Urbach, and Tilo Wettig.

I am grateful to Hubert Simma for valuable comments and suggestions to this contribution. The work of CS is supported by Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) - Project No. 315477589 (CRC-TR 211) and 460248186 (PUNCH4NFDI) as well by the competence network HPC.nrw of the state of North Rhine-Westphalia.

## References

- [1] UKQCD collaboration, *International lattice data grid*, *Nucl. Phys. B Proc. Suppl.* **119** (2003) 225 [[hep-lat/0209121](#)].
- [2] A.C. Irving, R.D. Kenway, C.M. Maynard and T. Yoshie, *Progress in building an International Lattice Data Grid*, *Nucl. Phys. B Proc. Suppl.* **129** (2004) 159 [[hep-lat/0309029](#)].
- [3] A. Ukawa, *Status of international lattice data grid: An Overview*, *Nucl. Phys. B Proc. Suppl.* **140** (2005) 207 [[hep-lat/0409084](#)].

- [4] M.G. Beckett, B. Joo, C.M. Maynard, D. Pleiter, O. Tatebe and T. Yoshie, *Building the International Lattice Data Grid*, *Comput. Phys. Commun.* **182** (2011) 1208 [0910.1692].
- [5] ILDG, “International Lattice Data Grid.” <https://hpc.desy.de/ildg/organization/>.
- [6] ILDG, “Identity and Access Management (IAM).” <https://iam-ildg.cloud.cnaf.infn.it/>.
- [7] CSSM, “Centre for the Subatomic Structure of Matter.” <https://set.adelaide.edu.au/physics-chemistry-earth-sciences/our-research/physics/nuclear-and-particle-physics/cssm>.
- [8] JLDG, “Japan Lattice Data Grid.” <https://www.jldg.org/>.
- [9] LDG, “LatFor Data Grid.” <https://hpc.desy.de/ldg/>.
- [10] UKLFT, “UK Lattice Field Theory.” <http://generic.wordpress.soton.ac.uk/uklft/>.
- [11] USQCD, “US Lattice Quantum Chromodynamics.” <https://www.usqcd.org/>.
- [12] ILDG, “VO policy.” [https://hpc.desy.de/ildg/vo\\_policy/](https://hpc.desy.de/ildg/vo_policy/).
- [13] ILDG, “Acceptable Use Policy.” [https://hpc.desy.de/ildg\\_webdocs/ILDG-AUP-v0.2-20240701.html](https://hpc.desy.de/ildg_webdocs/ILDG-AUP-v0.2-20240701.html).
- [14] M.D. Wilkinson, M. Dumontier, I.J. Aalbersberg, G. Appleton, M. Axton, A. Baak et al., *The fair guiding principles for scientific data management and stewardship*, *Scientific Data* **3** (2016) 160018.
- [15] ILDG, “Hands-on material.” <https://gitlab.desy.de/ildg/hands-on>.
- [16] curl. <https://curl.se/>.
- [17] ILDG, “Swagger documentation of the RESTful API of the metadata catalogue.” <https://idefix-vm10.zeuthen.desy.de/ildg/mdc/swagger-ui/index.html>.
- [18] D. Pleiter and H. Simma, “Simple listings.” <https://www-zeuthen.desy.de/ape-cgi-bin/ildg-mdc.cgi>.
- [19] H. Matsufuru, *QCDml markup tool*, 2023. <https://www.jldg.org/QCDml/>.
- [20] G. Pederiva, *Search and markup GUI*, 2025. <https://mdc.zam.kfa-juelich.de/search/ensemble>.
- [21] G. Bali et al., *Lattice gauge ensembles and data management*, *PoS LATTICE2022* (2022) 203 [2212.10138].
- [22] Y. Aoki et al., *Lattice gauge ensembles and data management*, *PoS LATTICE2024* (2025) 412 [2502.08303].
- [23] ILDG, “Hands-on workshop.” <https://indico.desy.de/event/49335>.
- [24] “Announcements for the european lattice field theory community.” <https://lists.geant.org/sympa/info/eurolft-announce>.