Integration of the NCRC Database and Other INL Databases

INL/RPT-24-79386

Nuclear Computational Resource Center

JULY 2024

Jason Weninger and Garrick Evans

AI, Data, and Visualization, Idaho National Laboratory

Jason Hales

Advanced Scientific Computing, Idaho National Laboratory

Idaho National Laboratory

Battelle Energy Alliance manages INL for the U.S. Department of Energy's Office of Nuclear Energy

INL/RPT-24-79386

Integration of the NCRC Database and Other INL Databases

Jason Weninger and
Garrick Evans
AI, Data, and Visualization, Idaho National Laboratory
Jason Hales
Advanced Scientific Computing, Idaho National Laboratory

July 2024

Idaho National Laboratory

Advanced Scientific Computing

Idaho Falls, Idaho 83415

inl.gov

Prepared for the
U.S. Department of Energy
Office of Nuclear Energy
Under DOE Idaho Operations Office
Contract DE-AC07-05ID14517

Page intentionally left blank

ABSTRACT

The Nuclear Computational Resource Center provides a portal by which industry professionals, educational staff, students, national laboratory employees, and others may request access to certain engineering software tools. As the tools provided through the Nuclear Computational Resource Center portal are not open-source and freely available, a set of approvals are necessary before access is granted. All code recipients must be associated with an institution that has a license with Idaho National Laboratory for the code requested. Information about these licenses is controlled by Idaho National Laboratory's Technology Deployment organization and housed in a Technology Deployment database. Those requesting code access who are not citizens of the United States must also have a security plan, mandated by Idaho National Laboratory policy. Security plans are managed by the International Access Program and are stored in an International Access Program database known as IFacts. Granting access to software thus depends on information stored in the Technology Deployment database and IFacts. In the past, no connection between the Nuclear Computational Resource Center portal and these databases existed, making checking the status of license agreements and security plans time consuming and error prone. This report demonstrates that the Nuclear Computational Resource Center portal now connects to both the Technology Deployment database and IFacts, greatly improving the ease of use of the Nuclear Computational Resource Center system by administrators, which leads to a better overall experience for those requesting code access.

ACKNOWLEDGMENTS

Rob Podgorney led the NCRC team until May 2024, and we appreciate his contributions. Others who made meaningful contributions include: Wendy Skinner, Business Analyst & Relations Specialist in the Technology Deployment Organization, who manages the IPData database and assisted in getting Technology Deployment information into NCRC; Raymond George, Software Developer in the Application Services Organization, a developer of the IFacts website; David Othus, Database Administrator in the Database Services Organization, who set permissions for HPC to access the IPData database; Alex Pospical, Software Developer in the Cloud & Platform Administration Organization, who created the NCRC Azure Storage account, connection strings, and security tokens; Keith Banner, Software Developer in the Application Development & Database Services Organization, who granted NCRC developers access to INL NuGet packages, specifically the INL Storage NuGet Package; Casey Killpack, Software Developer in the Dev Sec Ops Organization, who added tokens into the Azure DevOps vault; and Brett Hiltbrand, Quality Engineer in the Dev Sec Ops Organization, who managed the deployment of the NCRC portal to the production environment of the DMZ.

Page intentionally left blank

CONTENTS

FIGURES

ACRONYMS

COG	Code Oversight Group
DMZ	demilitarized zone
DOE	United States Department of Energy
HPC	High Performance Computing
IAP	International Access Program
INL	Idaho National Laboratory
MOOSE	Multiphysics Object-Oriented Simulation Environment
NCRC	Nuclear Computational Resource Center
NEAMS	Nuclear Energy Advanced Modeling and Simulation
SQL	Structured Query Language
TD	Technology Deployment

Integration of the NCRC Database and Other INL Databases

1 INTRODUCTION

The Nuclear Computational Resource Center (NCRC) application is an external-facing service established at Idaho National Laboratory (INL) to simplify requests to access computational tools, High Performance Computing (HPC) resources, and training (see Figure 1, inl.gov/ncrc). The NCRC portal is funded by the United States Department of Energy (DOE)'s Nuclear Energy Office, which also funds the development of a suite of nuclear engineering codes through the Nuclear Energy Advanced Modeling and Simulation (NEAMS) program and based on the Multiphysics Object-Oriented Simulation Environment (MOOSE) modeling and simulation framework. The Office of Nuclear Energy recognizes the need for an efficient process for distributing these engineering codes to users and giving them the computing resources to take full advantage of the codes.

The NCRC portal target audience includes, but is not limited to, users from the private sector, national laboratories, universities, and federal agencies. Available computational tools include Bison, a nuclear fuel performance code [1]; Griffin, a tool for neutronics calculations [2]; Pronghorn, a coarse-mesh, thermal-hydraulics code [3]; and others. Descriptions of these codes are on a page on the NCRC website.

The NCRC portal is also a tool for NCRC administrators to gather, store, and manage personal data and to track users that have requested access (or already have access) to HPC accounts and software.

The NCRC application resides in INL's so-called demilitarized zone (DMZ). The DMZ is a perimeter network that protects and adds an extra layer of security to INL's internal network from untrusted traffic. The goal of the DMZ is to allow INL access to untrusted networks, such as the internet, while ensuring INL's internal network remains secure

Inside INL's internal network, but outside the DMZ, are two databases with information critical to the NCRC workflow. Currently, NCRC administrators are using email to transfer data between the NCRC application and these databases. The first of these databases is IPData, a database of software license information. IPData is maintained by the TD group. More than 750 license agreements associated with NCRC software are recorded in the IPData database.

Before users obtain access to licensed software, a license agreement must be in place between the user's institution and INL. IPData stores this information, which must be accessed manually by NCRC personnel to verify that a given user is covered under an active license. Until this project, no connection existed between the IPData and NCRC databases.

The second database needed by NCRC is the IFacts database managed by the IAP group. This database houses information about international employees and international collaborators. The information is stored in security plans, and each international employee or collaborator with some level of access at INL has an

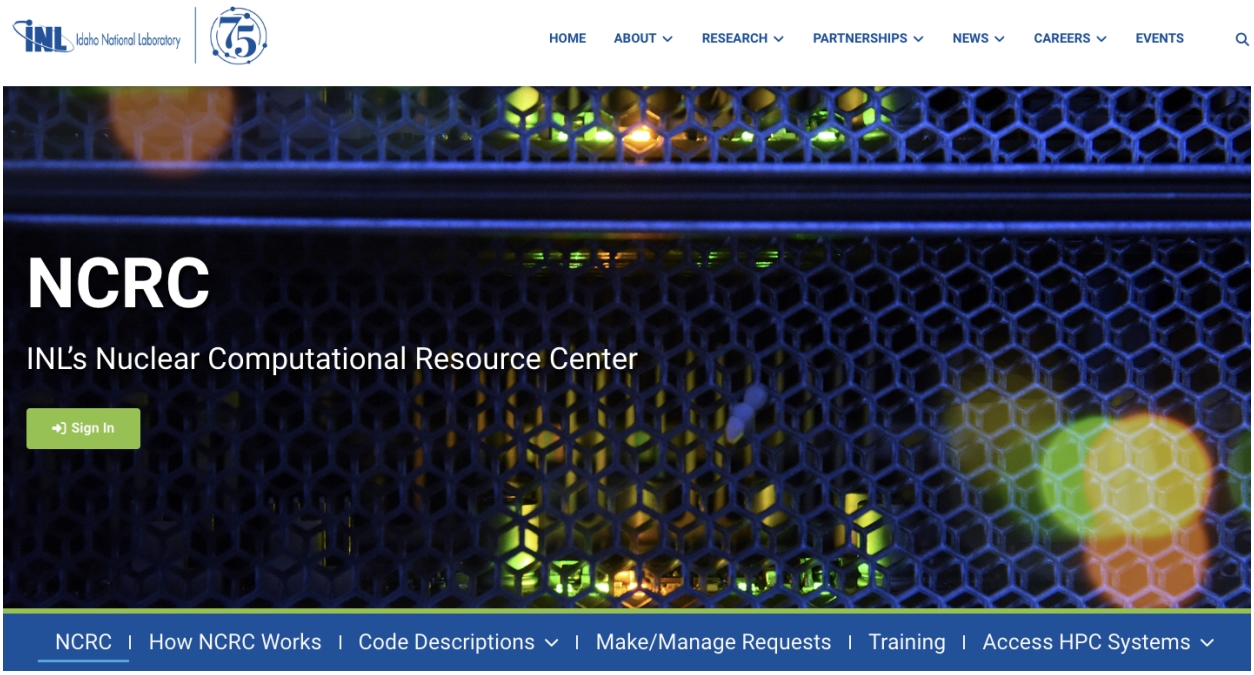

Figure 1. The NCRC home page.

active security plan. This database includes nearly 1,000 records of active security plans alone.

Before international staff or collaborators may receive access to export controlled or licensed software through the NCRC portal, they must have an INL security plan. As with the IPData database, retrieving information from the IFacts database is a manual step for NCRC personnel, with no connection between the IFacts and NCRC databases.

The lack of connection between the NCRC, IPData, and IFacts databases places an unnecessary burden on NCRC staff since a manual check must be made to find information in the IPData and IFacts databases to verify that the proper agreements and documentation are in place. This lack of connection introduces the opportunity for human error as these databases are checked. Finally, this lack of connection slows the approval process, a frustration for NCRC staff and users alike. All of these drawbacks result in higher administrative costs for maintaining the NCRC database and processing user requests.

The purpose of this project is to enable a connection between the NCRC database and the IPData and IFacts databases. This report continues with a discussion of Azure Storage. The IFacts database and the connection made between the NCRC and IFacts databases are explained. A similar discussion focused on the

IPData database then follows. Next is a review of the new functionality provided by this project, which is followed by a discussion of future work, additional achievements in FY-24 beyond the scope of the database work, and a summary.

2 BACKGROUND INFORMATION ABOUT AZURE STORAGE

In FY-23 NCRC staff held several planning meetings with multiple organizations to discuss establishing a connection between the NCRC application and databases within INL's internal network. Azure Storage was approved as a secure way to transfer data between NCRC in the DMZ and INL's internal network databases.

The first iteration of NCRC interfacing with INL's internal network databases involved using Azure Queue Storage. Azure Queue Storage is a service for storing messages that can be accessed via authenticated HTTPS calls. This process relied on other organizations to create their own .NET Windows service to check the Azure Queue Storage for messages sent from NCRC. The internal service would then search the appropriate database for the data requested and send a message back to the Queue with the correct data. The NCRC application would need to check the Queue periodically for any response. While testing this concept, the process proved to be slow and inefficient. Another drawback was that it relied on other organizations to create and maintain a service NCRC required.

NCRC is now interfacing with INL's internal network databases using Azure Storage Tables via HPC. Specific data is gathered from internal network databases and transferred to Azure Storage Tables. This allows NCRC staff to query data in near real time. Details are in the sections that follow. To complete the software changes that allow this data transfer, NCRC developers needed to coordinate with several people in multiple organizations. This coordination was a critical piece of the success of the project. We appreciate the timely help received from many.

3 CONNECTION TO IFACTS

To begin the discussion of the IAP database, IFacts, consider Figure 2. In Figure 2, the INL network is shown as consisting of two parts: the DMZ and the internal network. The NCRC application resides in the DMZ. The IFacts database resides in the INL internal network. With a firewall between the DMZ and the INL internal network, it is impossible for the NCRC application to communicate with the IFacts database directly.

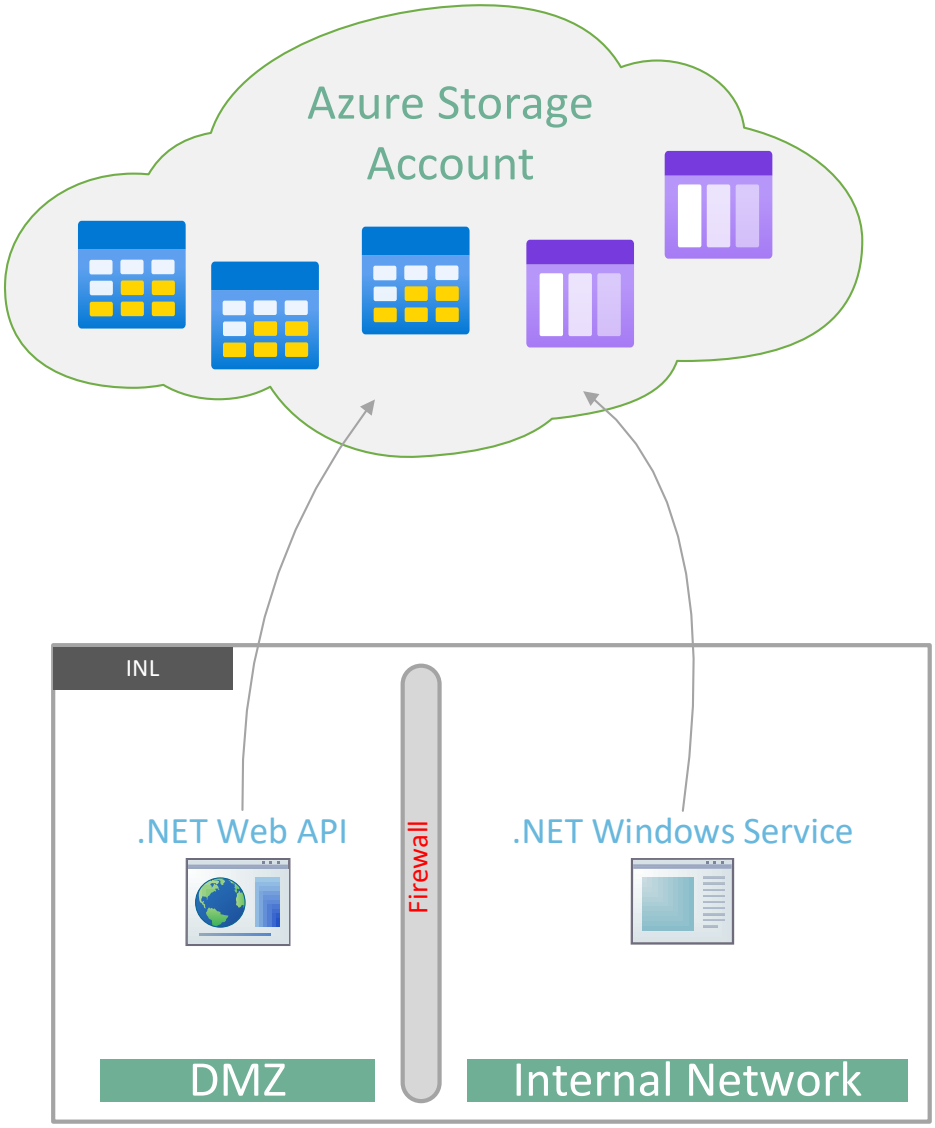

Figure 2. Schematic of DMZ, INL internal network, and Azure storage locations.

INL's HPC maintains its own database of information associated with HPC accounts. This system is located in the INL internal network (not in the DMZ). The HPC team had previously configured a connection to the IFacts database to bring needed data into the HPC database. This was necessary for HPC's own purposes.

The existing pipeline involved a Python script invoked by a cron job. Cron is a job scheduler that runs jobs at fixed intervals. The Python script connects to the Structured Query Language (SQL) Server that stores the IFacts tables. The script then collects the pertinent IFacts data. Next, the script connects to HPC's Postgres database and overwrites all IFacts data there, ensuring HPC has up-to-date IFacts information.

Leveraging that existing connection and data transfer, an automated script now transfers the IFacts data from the HPC database into the Azure storage space. This is done by another Python script on a scheduled cron job. This script queries the HPC Postgres database and converts the IFacts data into an HTTPS POST request that will insert the IFacts records into the Azure storage space. The Azure storage space can then be accessed by the NCRC application at any time.

This new connection has several advantages. NCRC staff now have a direct connection to IFacts data, with the necessary information appearing in one place. This speeds the process of account approval and reduces the opportunity for error. This change also enables additional enhancements, to be discussed later.

It is worthwhile to note that the processing and movement of this data is performed according to INL's security protocols, protecting both the INL network and users' sensitive information.

4 CONNECTION TO IPDATA

The pattern built for the IFacts database connection was followed for the needed connection between the NCRC and IPData databases.

First, a new connection needed to be made to reach the SQL server that the IPData database exists on. Permission to access the IPData database was given by the Technology Deployment Organization. A database administrator in the Database Services Organization was able to create a login we could use to connect to IPData, and HPC created a new table within their database to hold IPData's information. Next, a scheduled Python script connects to the IPData SQL server and copies the data into HPC's database. Finally, another Python script translates the data from the Postgres database and sends each record to the Azure storage space as an HTTPS POST request, just like the IFacts connection. Once in Azure storage, the NCRC application may access the data at any time.

This connection brings licensing information into the NCRC system. Staff may now review license agreement status for the user's home institution from within NCRC. This simplifies the process of verifying a user's license status for export controlled and licensed simulation software requests.

5 NEW FUNCTIONALITY

From the perspective of NCRC users, the changes discussed above have no effect on the NCRC application. For NCRC staff, two new views are available. See Figure 3, which shows links to two new dashboards.

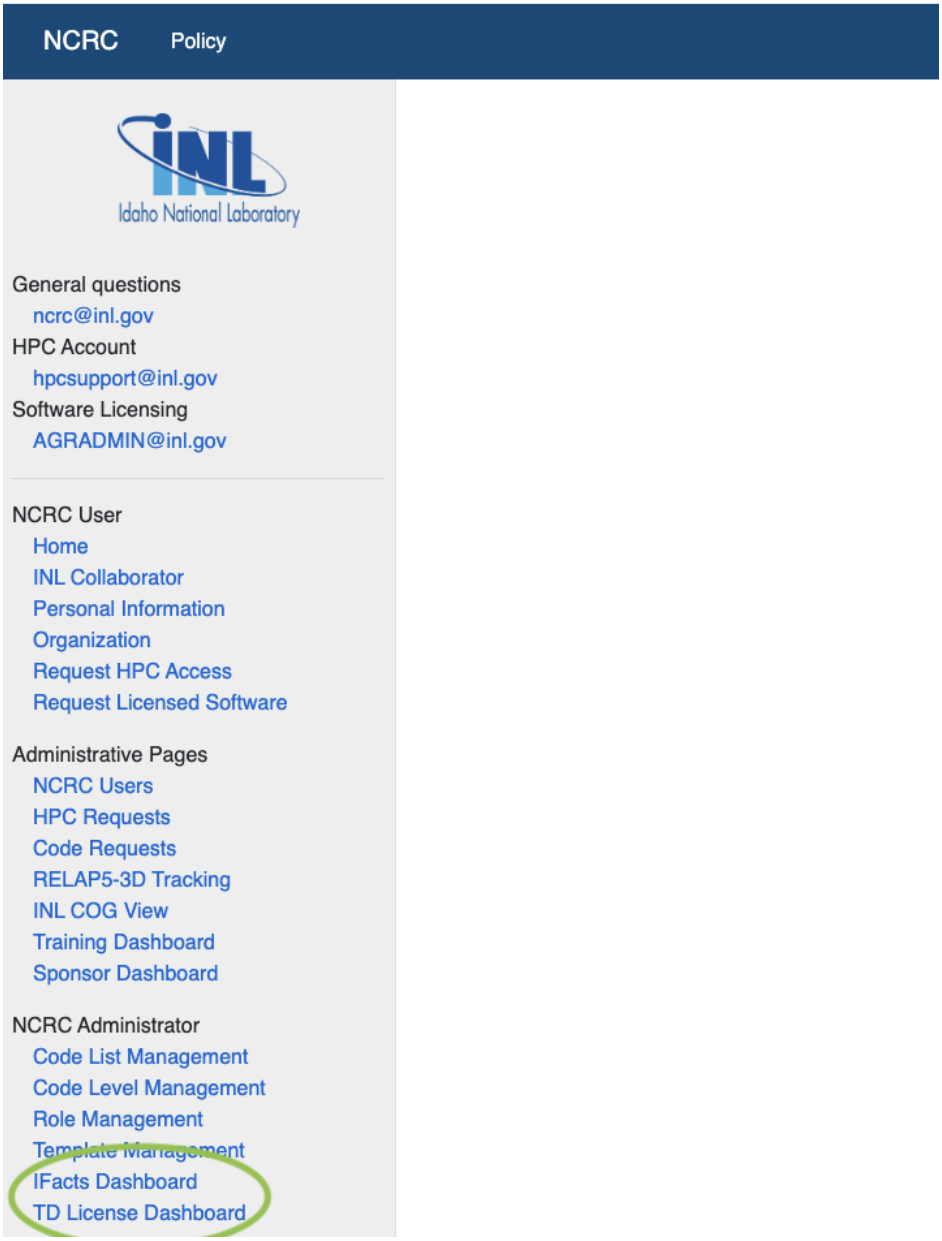

Figure 3. New dashboards available to NCRC staff.

The first link is associated with INL security plans (see Figure 4). On this page, NCRC staff can search for security plans associated with any NCRC user.

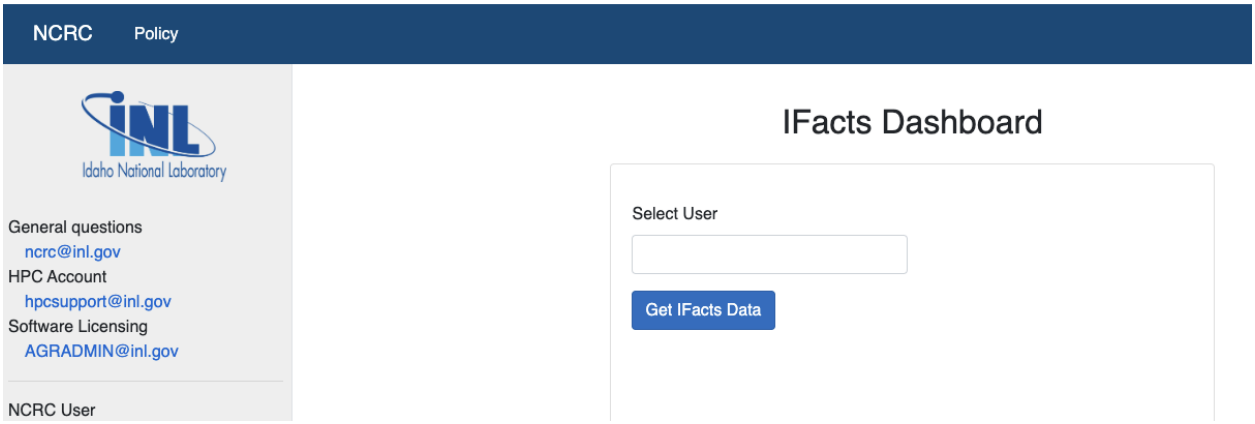

Figure 4. New interface for searching for security plans.

Results of a query are shown in Figure 5. All security plans associated with a given user are displayed. It is possible in some cases for a user to have an active or complete plan and another plan in the the development stage. Seeing all plans at once helps NCRC staff quickly navigate to the appropriate plan for the task or question at hand.

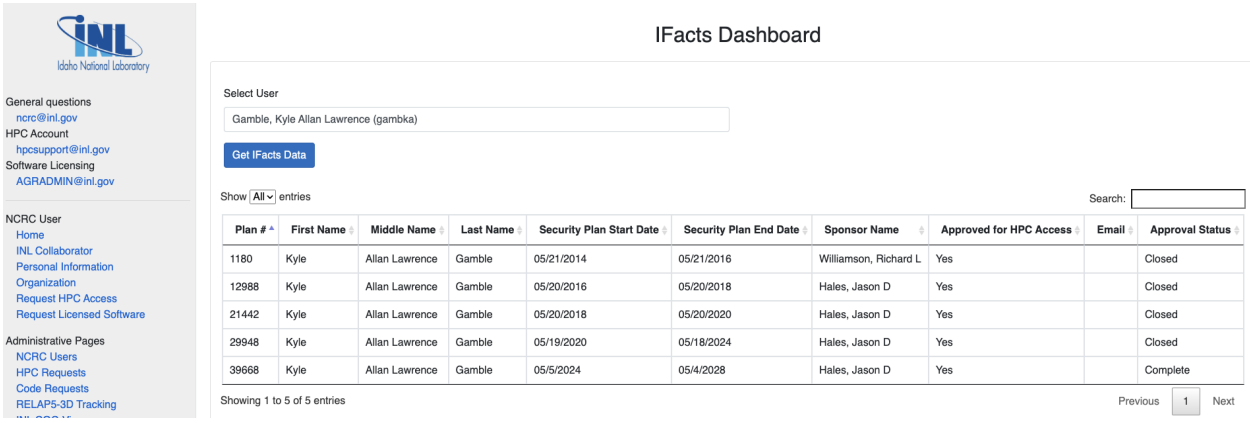

Figure 5. Results from a sample query of security plans.

The TD License Dashboard link shows a list of all license agreements associated with NCRC. See Figure 6. By entering text into the search box at the top right of the page, NCRC staff can find all agreements associated with a given company (e.g., Westinghouse) or for a particular software application, (e.g., Bison). This greatly simplifies checking whether an institution has an active license agreement for a given software application.

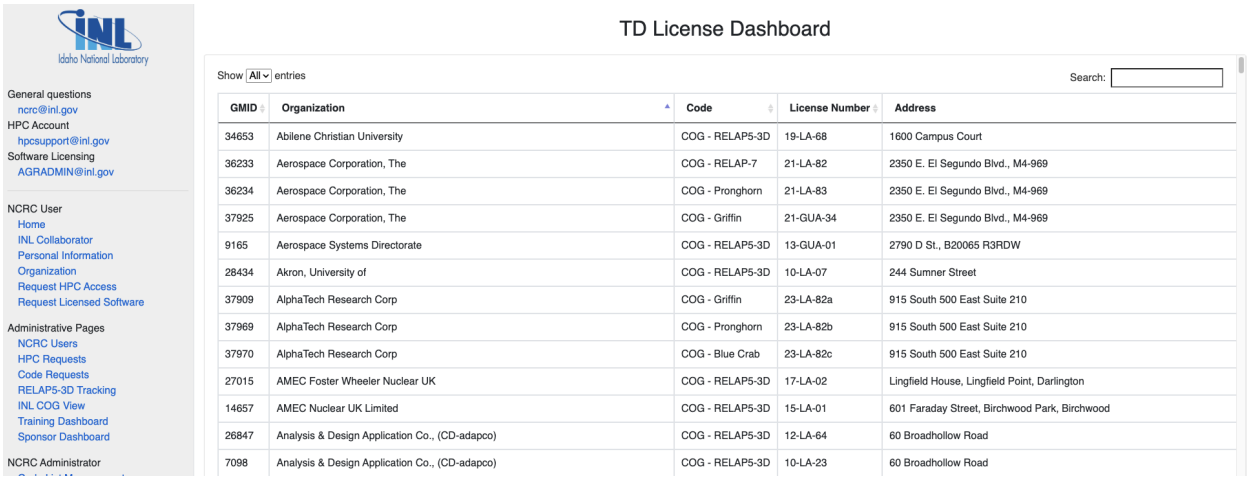

Figure 6. New Technology Deployment license database view.

6 ENABLED FUTURE WORK

The connections described above are foundational for a number of potential enhancements to the NCRC portal:

- Since the IFacts data is accessible from within the NCRC portal, it would be possible to provide and revoke HPC and code access on a day-to-day basis, based on a changed status in a security plan.
- The plan numbers returned as part of a query of a user on the IFacts page from within the NCRC portal (see Figure 5) can be made to be links to the specific plan page on the IAP website.
- Filters can be added to the set of security plans returned such that only the active plan (for example) is displayed.
- Information about code requests and approvals as stored in the IFacts database could be returned to the NCRC portal. For example, all software package requests fully approved in the IFacts database could be shown.
- For security plans in a pending state, the list of outstanding approvals could be provided, along with the elapsed time since the request for approval was submitted. This would help NCRC staff quickly know if an approval request has been forgotten and who needs a reminder.
- With assistance from the IFacts team, security plan data originating in the NCRC portal could be pushed to the IFacts database, accelerating the process for creating new security plans.
- Since the IPData data is accessible from within the NCRC portal, it would be possible to provide and revoke code access on a day-to-day basis, based on a changed status in a license agreement.
- Filters can be added to the set of license agreements on the TD License Dashboard page such that only the active agreements (for example) are displayed.
- With assistance from the IPData team, license data originating in the NCRC portal could be pushed to the IPData database, accelerating the process for creating new license agreements.

7 ADDITIONAL ACHIEVEMENTS

A number of additional enhancements to the NCRC portal have been under development in FY-24. These have in common the need to improve the user experience, both of HPC and code request users and NCRC administrators. These improvements include:

- Rework of NCRC database: Creating new tables to make a more relational database. User information, organization information, HPC requests, code level requests, and justifications will now be in separate tables.
- Complete Rework of Email system: Each page had its own code to create and send emails. An Email helper is being created to eliminate duplicated code.
- Implementation of new questions required by Technical Deployment: Code justification questions are being modified and will require more detailed answers. Several new questions are being added for NPAC (Nonproliferation and Arms Control) or 810 codes.
- Organization Licensed Point of Contact Dashboard: This dashboard will give the point of contact of an organization the ability to see all users in their organization who have requested code, a view of the status of the request, and the ability to approve or deny code access on a per-user basis.
- Admin review Dashboard: This dashboard will allow NCRC administrators to verify the user's organization and to review the user's uploaded documents. This will be one of the places administrators can return a request to the user if more information is needed before HPC and Code Oversight Group (COG) reviews.
- IAP Verification Dashboard: This dashboard will allow administrators to see all international collaborators who have made a request for code or an HPC account. The administrator will be able to see if the user has a security plan number, search IFacts for a security plan, and add security plan numbers and expiration dates to the NCRC database.
- Added reCAPTCHA to registration and training registration: This was added to help prevent the creation of accounts by bots.
- Enabling HPC requests and code requests to have an independent workflow and allowing users to request multiple levels of code access: These two connected items have required a rework of the entire NCRC portal. Today, an HPC account request must be made before a user can request code. With these changes, users may request code without requesting an HPC account as a prerequisite step. Additionally, certain code requests will be reviewed and approved independently of the user having an active or requested HPC account.

These improvements are nearly ready for release and are the current focus of the NCRC development and maintenance team. It is expected that most of these changes, if not all, will be finalized over the next month.

Following these changes, the NCRC team will assess the current state of the portal and prioritize the next set of improvements (many of which are given in Section 6). These priorities will form a basis for activities in FY-25.

8 SUMMARY

The NCRC portal is the entry point by which industry professionals, educational staff, students, national laboratory employees, and others may request access to licensable, export controlled nuclear energy software maintained by INL. As the tools provided through the Nuclear Computational Resource Center portal are not open-source and freely available, a set of approvals are necessary before access is granted. The first requirement is that code recipients be associated with an institution that has a license with Idaho National Laboratory for the code requested. INL's TD organization maintains information about these licenses in a TD database called IPData. Non-U.S. citizens requesting access must also have an INL security plan. The IAP organization maintains information associated with security plans. This information is in the IFacts database. Approval for access to NCRC software is predicated on appropriate information in the IPData and IFacts databases. Until this project, no connection between the Nuclear Computational Resource Center portal and these databases existed. NCRC staff were required to check the status of license agreements and security plans by copying information from the NCRC portal to the interface of other databases, a time consuming and inaccurate exercise. This report demonstrates that the Nuclear Computational Resource Center portal now connects to both the IPData and IFacts databases, greatly improving the ease of use of the Nuclear Computational Resource Center system for administrators. Other improvements are possible, which will further streamline the workflow for NCRC staff, leading to faster response times and a better overall experience for those requesting code access. These enhancements include links to specific IAP security plan pages, filters to further refine search results, and the ability to provide and revoke access more automatically based on the status of security plans and license agreements. Additional features have been identified that will further improve the user experience for NCRC staff and those making HPC and software requests. These will be prioritized and implemented during the remainder of FY-24 and in FY-25.

9 REFERENCES

[1] R. L. Williamson, J. D. Hales, S. R. Novascone, G. Pastore, K. A. Gamble, B. W. Spencer, W. Jiang, S. A. Pitts, A. Casagranda, D. Schwen, A. X. Zabriskie, A. Toptan, R. Gardner, C. Matthews, W. Liu, and H. Chen, "BISON: A flexible code for advanced simulation of the performance of multiple nuclear fuel forms," *Nuclear Technology*, vol. 207, no. 7, pp. 954–980, 2021.

[2] Y. Wang, Z. M. Prince, O. W. Calvin, H. Park, N. Choi, Y. S. Jung, S. Schunert, S. Kumar, J. T. Hanophy, V. M. Labouré, C. Lee, J. Ortensi, L. H. Harbour, and J. R. Harter, "Griffin: A MOOSE-based reactor physics application for multiphysics simulation of advanced nuclear reactors," *Annals of Nuclear Energy*, May 2024. Preprint submitted.

[3] A. Lindsay, G. Giudicelli, P. German, J. Peterson, Y. Wang, R. Freile, D. Andrs, P. Balestra, M. Tano, R. Hu, L. Zou, D. Gaston, C. Permann, and S. Schunert, "MOOSE Navier–Stokes module," *SoftwareX*, vol. 23, p. 101503, 2023.