

Nutzungsbedingungen und Datenschutzerklärung für die HPC-Systeme am ZIH

Stand: 1. April 2025

Die Person, die den Antrag stellt, erkennt die "Nutzungsbedingungen und Datenschutzerklärung für die HPC-Systeme am ZIH" an. In diesen Bedingungen ist der Umgang mit User-Daten auf den Hochleistungsrechnern geregelt. Die Bedingungen können von den Regelungen der IT-Ordnung¹ und den allgemeinen Benutzungsbedingungen des ZIH abweichen.

Die Hochleistungsrechner des ZIH sind hochspezialisierte Forschungsgeräte, die durch das ZIH für verschiedene User-Gruppen betrieben werden. Um einen störungsarmen Betrieb und die effiziente Wartung der Systeme zu gewährleisten, werden nachfolgende Bedingungen anerkannt.

§1 - Speicherung persönlicher Daten im HPC-Projektesystem

Das HPC-Projektesystem unterstützt die Angestellten des HPC-Supports bei Aufgaben des laufenden Betriebs (z.B. Projektmanagement, User-Betreuung, Statistiken). Hier werden neben den projektbezogenen Daten auch persönliche Kontaktdaten der HPC-User (Name, Vorname, Titel, dienstl. E-Mail-Adresse, dienstl. Telefonnummer, dienstl. Postanschrift) gespeichert. Der HPC-Support kann diese Kommunikationswege für folgende Zwecke benutzen:

- postalische Zusendung vertraulicher Informationen (z.B. Zugangsdaten),
- Benachrichtigung bei Ausfällen, Wartungen oder Schulungen per E-Mail,
- Anfrage des Supports bei User-spezifischen Problemen im Betrieb (z.B. Beeinträchtigung anderer User) per E-Mail oder Telefon.

Die Frist zur Aufbewahrung dieser Daten im HPC-Projektesystem richtet sich nach §4 (4) der „Ordnung zur Errichtung und zum Betrieb eines Identitätsmanagementsystems an der Technischen Universität Dresden“.² Nach Ablauf dieser Frist werden die personenbezogenen Daten der User in diesem System anonymisiert.

Ausgenommen hiervon sind eine Beschreibung des Projektes, der Nachweis der genutzten Ressourcen (Accounting) sowie die Kontaktdaten der Projektleitung. Diese werden dauerhaft für wissenschaftliche Zwecke und zur Gewährleistung einer effektiven Nutzung von Hochleistungsrechnern im Verbund für Nationales Hochleistungsrechnen (NHR) gespeichert und genutzt.

§2 - Weitergabe von persönlichen Daten

Für die deutschlandweite Koordinierung der Ressourcenvergabe im HPC-Bereich werden im Zuge der Projektbeantragung neben den Projektdaten für die wissenschaftliche Begutachtung die dienstlichen Kontaktdaten der beantragenden Projektleitung und der beteiligten Wissenschaftler an die Geschäftsstelle für Nationales Hochleistungsrechnen im DFN-Verein, Rosenstraße 2, 10178 Berlin weitergeleitet.

§3 - Arbeit im Projekt

Die Nutzung der HPC-Systeme ist ausschließlich für Forschungszwecke und im Rahmen des beantragten Projektes zulässig. Eine darüber hinausgehende oder abweichende Nutzung ist unzulässig.

Die Gesamtverantwortung (*Principal Investigator*) für alle Aktivitäten innerhalb des Projektes auf den HPC-Systemen des ZIH hat in der Regel eine Leiterin oder ein Leiter einer Forschungsgruppe (z.B. Lehrstuhlinhaber/in). Es kann eine weitere Person (*Projektadmin*) benannt werden, die für die technisch-administrativen Belange des HPC-Projektes zuständig ist und dem ZIH als direkter Ansprechpartner dient. Beide werden im Weiteren als Projektleitung bezeichnet.

Die Projektleitung ist insbesondere befugt:

- User hinzuzufügen - dabei sind u.U. spezielle gesetzliche Regelungen für ausländische Bürger zu berücksichtigen;
- Kontaktdaten der Mitglieder bei Änderung zu aktualisieren;
- HPC-Zugänge von Mitgliedern, die aus der Arbeitsgruppe ausgeschieden sind, zu sperren;
- die benötigten Ressourcen des Projektes zu überwachen;
- vor Ausscheiden von Usern deren Daten im Projektverzeichnis zu sichten und sicherungswürdige Daten zu speichern;
- User aus dem Projekt zu entfernen.

Für den effizienten Betrieb der HPC-Systeme, zur Übersicht und statistischen Auswertung werden Daten zur Ausnutzung von Ressourcen (CPU, RAM, Dateisystem, Software, Energie) jedes Jobs automatisch erfasst (Accounting). Dabei gilt folgendes:

- Die Projektleitung erhält Zugang zu erfassten Performance-Metriken für die Jobs der Projektmitglieder.
- Dargestellt wird der Ressourcenverbrauch der Jobs eines Users oder für das gesamte Projekt, um gegebenenfalls zwischen den Mitgliedern eines Projekts koordinieren zu können.
- Das ZIH nutzt gegenüber dem wissenschaftlichen Beirat des ZIH, dem CIO der TU Dresden und dessen Beirat oder anderen übergreifenden Gremien nur eine projektbezogene Accounting-Sicht.
- Innerhalb einer Frist gemäß §4 (4) der „Ordnung zur Errichtung und zum Betrieb eines Identitätsmanagementsystems an der Technischen Universität Dresden“ werden die zugehörigen personenbezogenen Accounting-Daten im Projektesystem anonymisiert.

Mit dem Ende der Laufzeit eines Projekts können keine neuen Jobs mehr gestartet werden. Zum Zwecke der Sicherung wichtiger Forschungsdaten können die Zugangsberechtigungen für die Mitglieder dieses Projekts noch weitere 30 Tage genutzt werden, bevor sie deaktiviert werden. 100 Tage nach Projektende werden die dem Projekt zugeordneten Daten in den HPC-Dateisystemen gelöscht.

§4 - Umgang mit Daten

Für die wissenschaftliche Arbeit auf den Hochleistungsrechnern werden verschiedene Dateisysteme mit unterschiedlichen Eigenschaften bereitgestellt. Für temporäre, häufig veränderliche Dateien und Verzeichnisse werden eigens für das Computing konzipierte Scratch-Dateisysteme zur Verfügung gestellt. Solche Dateien sollten nicht im /home oder /projects Dateisystem liegen, die für Job-Konfigurationen und die Dokumentation von Arbeitsabläufen und (Zwischen-) Ergebnissen vorgesehen sind. Für die langfristige Aufbewahrung von Daten können Archive am ZIH benutzt werden. Aus dieser Hierarchie der Speichersysteme im HPC-Bereich ergeben sich folgende Besonderheiten:

- Die HOME-Verzeichnisse der User unter /home sowie die Projektverzeichnisse unter /projects werden automatisch per Backup auf einem Bandsystem gesichert. Die Verweilfrist beträgt dort 180 Tage. Zusätzliche Snapshots innerhalb des Dateisystems erleichtern den Zugang zu älteren Dateien. Snapshots bleiben 3 Monate im System. Trotz Redundanz können Datenverluste nicht ausgeschlossen werden. Temporäre, häufig veränderliche Dateien und Verzeichnisse sollten dort nicht abgelegt werden.
- Für temporäre und häufig veränderliche Daten stehen Workspaces³ auf speziellen HPC-Dateisystemen zur Verfügung. Diese Workspaces werden nach Fristende automatisch gelöscht. Ergebnisse aus Berechnungen sollten entsprechend dem Lebenszyklus von Daten⁴ auf persistenten Dateisystemen gesichert oder archiviert werden.
- Daten von Knoten-lokalen Speichermedien werden in der Regel nach 7 Tagen (Login-Knoten) bzw. unmittelbar nach Job-Ende (Compute-Knoten) gelöscht.

Das ZIH geht davon aus, dass sämtliche Daten in den Dateisystemen der HPC-Maschinen projektbezogene Daten sind. Daraus ergeben sich die folgenden Regeln:

- Der Projektleitung kann Lese- und Schreibzugang zu den Projektverzeichnissen der Projektmitglieder gewährt werden. Dies gilt ausdrücklich auch für Verzeichnisse bereits ausgeschiedener User. Davon ausgenommen sind die HOME-Verzeichnisse.
- User und Projektleitung sind verpflichtet, *vor* dem Ausscheiden eines Users aus dessen Projekt die Projektdaten zu sichten und relevante Daten zu sichern. Dazu können z.B. die Speicherdienste des ZIH (Langzeitarchivierung von Forschungsdaten, Zwischenarchiv für unstrukturierte Daten) genutzt werden.
- 100 Tage nach dem Erlöschen der HPC-Zugangsberechtigung werden die User-Daten auf den HPC-Dateisystemen automatisch bereinigt. Daten im HOME-Verzeichnis des ehemaligen Users werden gelöscht. Die Daten des ehemaligen Users unter /projects werden der Projektleitung übertragen.

§5 - Probleme und Störungen im Betrieb

HPC-Systeme sind anfällig gegenüber Fehlern in der Nutzung. Der HPC-Support ist berechtigt, Jobs, die den normalen Betrieb beeinträchtigen, abzurechnen oder anzuhalten. Der User wird in diesem Fall informiert und in der Regel zur Fehlersuche aufgefordert. Der User verpflichtet sich, die Hinweise des HPC-Supports zu befolgen. Bei Zuwiderhandlung kann er von der Nutzung des Batch-Systems ausgeschlossen oder sein Login gesperrt werden.

Alle im HPC-Kontext anfallenden Log-Daten werden in einer Datenbank erfasst und für Langzeit-Analysen benutzt. Nach einem Monat werden Meldungen mit Login-Bezug gelöscht.

§6 - Erfassung von Performance-Daten zu User-Jobs

Für Forschungszwecke werden job- und login-bezogene Performance-Daten in einer Datenbank gespeichert und ausgewertet. Dazu gehören unter anderem Metriken aus den Prozessoren, Netzwerken, Speichersystemen und der Stromversorgung. User können ihre Job-Performance-Daten in einer Web-Ansicht⁵ einsehen, diese ist auch für die Projektleitung sichtbar. In anonymisierter Form können diese Daten an HPC-Forschungspartner weitergegeben werden.

Der HPC-Support erhält Zugang zu den Performance-Daten, um Optimierungspotential zu identifizieren und die User darüber zu informieren.

¹ <http://www.verw.tu-dresden.de/VerwRicht/Ordnung/download.asp?file=IT-Ordnung.pdf>

² <http://www.verw.tu-dresden.de/AmtBek/PDF-Dateien/2011-05/sonstO26.07.2011.pdf>

³ https://doc.zih.tu-dresden.de/data_lifecycle/working/

⁴ https://doc.zih.tu-dresden.de/data_lifecycle/overview/

⁵ <https://pika.zih.tu-dresden.de/>