

REGIMENTO INTERNO DA PLATAFORMA MULTIUSUÁRIO DE BIOINFORMÁTICA DO INCA

CAPÍTULO I – DAS DISPOSIÇÕES PRELIMINARES

Art. 1º – Da Natureza e Missão: A Plataforma Multiusuário de Bioinformática do Instituto Nacional de Câncer (INCA) é uma infraestrutura de computação de alto desempenho (HPC) concebida para prover e facilitar o acesso dos pesquisadores, alunos e colaboradores da Instituição aos recursos computacionais e de bioinformática. Sua missão é apoiar projetos científicos do Instituto Nacional de Câncer e seus parceiros que utilizam técnicas de bioinformática para análise de dados de biologia molecular, genômica, transcriptômica, proteômica e demais "ômicas". O INCA é responsável pelo desenvolvimento e coordenação de ações integradas na prevenção e controle do câncer no Brasil, incluindo pesquisa, ensino e assistência.

Art. 2º – Do Âmbito de Aplicação: O acesso à infraestrutura da Plataforma é oferecido a pesquisadores, alunos e funcionários do INCA, e seus colaboradores, para o desenvolvimento de projetos cadastrados por um pesquisador responsável. O uso da plataforma é centrado em projetos, exigindo que cada usuário esteja vinculado a, pelo menos, um projeto cadastrado.

Art. 3º – Da Infraestrutura e Software: A Plataforma dispõe de uma infraestrutura robusta, incluindo: um cluster de computação de alto desempenho (HPCC-Crab) com 40 unidades de processamento, um *fat node* com 3TB de RAM, 280 TB de armazenamento temporário e 500 TB exclusivos para dados de sequenciamento. O ambiente é gerenciado pelo *Slurm Workload Manager*, garantindo alocação eficiente de recursos. A plataforma também conta com servidores dedicados para análises em R/RStudio e acesso via Open OnDemand, ampliando a acessibilidade para os usuários.

CAPÍTULO II – DO CADASTRO E ACESSO

Art. 4º – Do Modelo Centrado em Projetos: O uso da Plataforma é centrado em projetos, o que significa que cada usuário deve estar vinculado a, pelo menos, um projeto cadastrado por um Pesquisador Principal / Investigador Responsável (PI) para solicitar sua conta. Esse modelo visa limitar o acesso aos dados apenas a usuários autorizados.

Art. 5º – Do Cadastro de Pesquisador Principal / Investigador Responsável (PI): Pesquisadores do INCA que desejam cadastrar projetos e usuários na Plataforma devem realizar, inicialmente, seu cadastro como Investigador Principal / Pesquisador Responsável. Nesta etapa, o PI deve informar seus dados (nome, e-mail institucional, telefone, ramal, laboratório) e concordar com os termos de responsabilidade e confidencialidade da Plataforma.

Art. 6º – Do Cadastro de Projeto: O cadastro de projeto é solicitado pelo PI. As informações necessárias incluem:

- Nome do Projeto (frase curta).
- Código identificador (8 a 16 caracteres, letras minúsculas, dígitos e traços). Este código será o nome do diretório e do grupo do projeto.
- Pesquisador Responsável.
- Descrição do projeto (tipo de dados, requisitos computacionais, espaço necessário).
- Usuários Participantes (nomes completos ou identificador do usuário).

Art. 7º – Do Cadastro de Usuário: O cadastro de usuários é solicitado pelo orientador (para alunos), ou pesquisador responsável (para colaboradores externos) ou pelo próprio pesquisador (pesquisadores do INCA).

- Os usuários precisam estar vinculados a projetos já cadastrados por um pesquisador.
- As informações do formulário de cadastro incluem:
 - Nome, Laboratório,
 - Ramal, e-mail, Telefones (celular e fixo),
 - Tipo de Usuário (PESQUISADOR, TECNOLOGISTA, ALUNO DO INCA ou COLABORADOR EXTERNO),
 - Pesquisador Responsável,
 - Login desejado (8 a 16 letras minúsculas)
 - Projetos (identificadores, separados por vírgula) aos quais o usuário estará vinculado.
- Usuários individuais recebem uma quota de 500 MB em seu diretório home.

Art. 8º – Do Cadastro de Pesquisadores de Outras Instituições: Para o cadastro de pesquisadores de outras instituições, é necessário que o requisitante envie uma solicitação formal para contatohpc@inca.gov.br, detalhando o projeto de pesquisa e a forma como os recursos da plataforma serão utilizados. Esta solicitação será submetida à avaliação do conselho gestor, que considerará as demandas computacionais, a relevância do projeto e a sua compatibilidade com as capacidades da Plataforma Multiusuário de Bioinformática.

Art. 9º – Do Primeiro Acesso e Alteração de Senha: No primeiro acesso, o usuário deve alterar a senha recebida por e-mail por uma senha segura.

Art. 10º – Da Inatividade de Contas: Contas de usuários que apresentarem inatividade superior a seis meses serão automaticamente suspensas. A reativação dessas contas poderá ser solicitada pelo pesquisador principal.

Art. 11º – Da Proibição de Compartilhamento de Credenciais: É proibido compartilhar dados de login de usuários.

CAPÍTULO III – DAS RESPONSABILIDADES

Art. 12º – Das Responsabilidades do Pesquisador Principal (PI): O Pesquisador Principal (PI) possui responsabilidades cruciais para a boa governança e utilização da Plataforma:

- Autorizar (e desautorizar) o acesso de usuários aos projetos.
- Notificar a Plataforma sobre o desligamento de qualquer usuário vinculado ao projeto.
- Definir os usuários com permissão de acesso ao diretório do projeto.
- Garantir que os recursos da plataforma sejam usados exclusivamente para o propósito do projeto cadastrado.
- Assegurar que os dados sejam transferidos para outra unidade de armazenamento (externo à plataforma) após a finalização dos projetos, visando a liberação de espaço.
- É o responsável pelo backup de seus dados de projeto (dados brutos, resultados, documentação, scripts e demais arquivos essenciais para reprodutibilidade).
- Orientar os usuários sobre o destino dos dados pós-processamento, observando as boas práticas de segurança e privacidade de dados pessoais, quando aplicável.
- Garantir a veracidade e precisão das informações fornecidas no cadastramento de projetos e usuários.
- Em caso de uso de dados genômicos de pacientes, o PI compromete-se a:
 - Obter o consentimento informado dos pacientes.
 - Assegurar que os dados sejam anonimizados ou codificados sempre que possível.
 - Adotar medidas rigorosas de segurança.
 - Utilizar os dados apenas para os fins especificados no consentimento e projeto aprovado.
 - Assumir a responsabilidade e prestar contas pelo uso ético, legal e transparente desses dados.

Art. 13º – Das Responsabilidades do Usuário: Todo usuário da Plataforma Multiusuário de Bioinformática do INCA compromete-se a:

- Concordar com o Termo de Responsabilidade e Confidencialidade da plataforma.
- Não divulgar sua conta e senha para qualquer pessoa.
- Respeitar as diretrizes de uso das redes às quais está conectado e as políticas da plataforma.
- Utilizar os recursos da plataforma exclusivamente para os propósitos explícitos do projeto cadastrado.
- Não utilizar o sistema para outra finalidade ou para gerar benefício próprio exclusivo e/ou unilateral.
- Não efetuar gravação ou cópia da documentação confidencial a que tiver acesso, nem apropriar-se de material confidencial e/ou sigiloso.
- Manter a confidencialidade dos dados acessados e não compartilhá-los com terceiros não autorizados.
- Tomar medidas adequadas para proteger os dados contra acesso não autorizado, incluindo o uso de senhas seguras.
- Fornecer consentimento explícito para o tratamento de seus dados pessoais pela plataforma.
- Seguir as melhores práticas de segurança da informação e estar ciente de que suas atividades podem ser monitoradas.

- Assumir a responsabilidade por quaisquer violações de dados causadas por negligência ou violação dos termos.
- Utilizar sempre navegação anônima (janela privativa) ao acessar o Open OnDemand para garantir o encerramento da sessão.
- Em caso de uso de dados genômicos de pacientes, o usuário concorda em:
 - Respeitar os termos do consentimento informado obtido dos pacientes.
 - Utilizar apenas dados anonimizados ou codificados.
 - Respeitar a confidencialidade e a privacidade.
 - Utilizar os dados exclusivamente para os propósitos do projeto.
 - Assumir a responsabilidade pelo uso adequado e relatar uso indevido.

Art. 14º – Das Responsabilidades da Equipe de Suporte da Plataforma: A Equipe de Suporte da Plataforma Multiusuário de Bioinformática é responsável por:

- Otimizar o atendimento às demandas dos usuários, registrando e monitorando solicitações.
- Fornecer suporte técnico e monitorar processos de dados de grande volume.
- Operar um sistema de tickets (OS-Ticket) local, acessível via VPN institucional, em <https://suporte-bioinfo.inca.gov.br/> para registro e acompanhamento de chamados.
- Comunicar proativamente sobre manutenções programadas e interrupções de serviço, orientando os usuários sobre o desligamento de jobs.
- Receber e avaliar sugestões para aprimorar as opções de suporte.
- Ser o canal de contato para dúvidas, preocupações ou reclamações relacionadas à proteção de dados.
- Informar a TI para liberação do protocolo de transferência de dados.
- Realizar o processo de transferência de dados seguindo as melhores práticas de segurança.

CAPÍTULO IV – DA GESTÃO DE DADOS E RECURSOS

Art. 15º – Das Cotas de Armazenamento: Cada Investigador Principal possui uma quota de 10 TB de armazenamento para seus projetos. A soma do armazenamento de todos os projetos vinculados a um PI não pode ultrapassar essa quota, salvo em casos especiais, autorizados pelo Conselho de Gestão da Plataforma. Em relação às cotas de armazenamento para os projetos de pesquisadores externos ao INCA, estes terão uma limitação de 2TB para dados em processamento. O prazo inicial para esta alocação é de 6 (seis) meses, com a possibilidade de extensão por mais 1 (um) ano mediante solicitação.

Art. 16º – Do Monitoramento de Armazenamento: Os usuários e PIs são solicitados a verificar periodicamente o volume de dados de seus projetos no espaço de armazenamento compartilhado. Dados de projetos finalizados ou não mais utilizados devem ser considerados para remoção, especialmente aqueles que estiverem há mais de 6 (seis) meses sem utilização. Ferramentas como o comando `du` podem ser usadas para este monitoramento.

Art. 17º – Da Remoção e Transferência de Dados

- Após a finalização de projetos, os dados devem ser removidos da Plataforma ou transferidos para outra unidade de armazenamento externo, para liberar espaço.
- A remoção de diretórios com grande número de arquivos deve ser feita com cautela, sendo recomendado o uso do comando `rsync` com a opção `--delete` para sincronizar um diretório vazio sobre o diretório a ser apagado, por ser mais eficiente que `rm`. É crucial lembrar que `rm` e `rsync` não utilizam lixeira, excluindo arquivos definitivamente. Recomenda-se o uso do `sleep 5` como medida de segurança antes de executar comandos de exclusão.

Art. 18º – Da Alocação de Recursos Computacionais: Sendo o HPCC-Crab um cluster multiusuário, as tarefas devem ser executadas exclusivamente nos nós de computação (*compute nodes*). Nenhuma análise pode ser executada no nó de login. O acesso aos recursos é gerenciado pelo *Slurm Workload Manager*, que permite submissão de tarefas via scripts ou acesso interativo a nós de computação.

Art. 19º – Das Ferramentas e Ambientes A Plataforma oferece diversas ferramentas e ambientes, incluindo:

- Módulos de Ambiente: Permitem simplificar a configuração do ambiente do usuário, carregar programas específicos sem afetar configurações padrão e criar módulos personalizados.
- CONDA Package Manager: Para instalação de programas e criação de ambientes virtuais Python.
- R-Studio Server: Acessível via web browser, para desenvolvimento e teste de scripts R. Análises pesadas devem ser submetidas via scripts Slurm.
- Open OnDemand: Um portal web para acesso integrado a todos os recursos HPC, incluindo navegador de arquivos, monitoramento de jobs, terminal BASH e aplicações interativas como Code Server, Jupyter Notebook e R-Studio Server.

CAPÍTULO V – DAS PUBLICAÇÕES

Art. 20º – Da Menção em Publicações: É um requisito que publicações e artigos decorrentes de trabalhos desenvolvidos na Plataforma Multiusuário de Bioinformática do INCA mencionem o uso da plataforma, preferencialmente com a denominação "*Bioinformatics Core Facility (INCA-RJ)*" em inglês.

CAPÍTULO VI – DAS DISPOSIÇÕES FINAIS

Art. 21º – Das Alterações no Regimento: Este Regimento Interno pode ser atualizado periodicamente para refletir mudanças na legislação ou nas políticas da Plataforma, sendo os usuários notificados sobre tais atualizações.

Art. 22º – Do Canal de Contato: Em caso de dúvidas, preocupações, solicitações de cadastro ou reclamações relacionadas à Plataforma ou proteção de dados, os usuários podem entrar em contato com a Equipe de Suporte da Plataforma Multiusuário de Bioinformática pelo e-mail contatohpc@inca.gov.br ou pelo sistema interno de suporte da plataforma <https://suporte-bioinfo.inca.gov.br/>.

Art. 23º – Dos Casos Omissos: Os casos omissos neste Regimento Interno serão analisados e deliberados pelo Conselho Gestor da Plataforma, em conformidade com as diretrizes do INCA e a legislação vigente.

Histórico de Revisões:

Data: 29/08/2025

Versão: 1.0

Data: 02/09/2025

Versão: 1.1