

Charte d'utilisation des infrastructures de la plate-forme bio-informatique GenoToul

Table des matières

TABLE DES MATIERES	1
A) PREAMBULE.....	2
B) OBJECTIFS	2
C) CONDITIONS D'ACCES A LA PLATE-FORME	3
1) ACCES A L'INFRA + DEMANDES EXCEPTIONNELLES	3
2) SERVICES LIES AU COMPTE UTILISATEUR	3
3) ENGAGEMENTS DE LA PLATE-FORME	4
4) DUREE DE VALIDITE D'UN COMPTE/CONDITIONS DE FERMETURE	5
a) <i>Durée de validité par défaut d'un compte</i>	<i>5</i>
b) <i>Durée de validité d'un mot de passe.....</i>	<i>5</i>
c) <i>Fermeture d'un compte</i>	<i>5</i>
d) <i>Prolongation de compte</i>	<i>5</i>
5) SUPERVISION DU FONCTIONNEMENT	6
D) DEVOIRS DE L'UTILISATEUR.....	6
1) SOBRIETE NUMERIQUE	6
2) SOUMISSION DE JOBS DE CALCUL.....	6
3) STOCKAGE DES DONNEES/SAUVEGARDE.....	6
4) UTILISATION DES LOGICIELS/BANQUES.....	7
5) INFORMATIONS PERSONNELLES ET MOT DE PASSE	7
6) RESPECT DES DIRECTIVES	7
7) RESPECT DU CARACTERE CONFIDENTIEL.....	7
8) DONNEES A CARACTERE PERSONNEL.....	8
9) HEBERGEMENT DE MACHINES VIRTUELLES	8
E) PROPRIETE ET VALORISATION DES RESULTATS	8

A) Préambule

Historique des mises à jours :

- 09/2023 : Meilleure spécification de la durée des comptes des personnels du privé, règle concernant l'acceptation des comptes des personnes travaillant à l'étranger, nécessité de récupérer les données d'un compte d'un collaborateur avant sa fermeture, correction d'erreur typographique, ajout de sobriété numérique et du PGD.
- 06/2019 : Mise en conformité RGPD et nouvelle politique de gestion des comptes + changement mise en forme
- 07/2017 : Rajout des infos sur la propriété des résultats
- 01/2016 : Précision droits des répertoires par défaut
- Précision respect des licences logicielles concernant aussi les utilisateurs de Galaxy
- 03/2015 : Interdiction de lancer tout traitement de données sur les serveurs frontaux et purge du work sans préavis si fichiers +120j
- 06/2014 : Création du document

Documents rattachés :

- [Charte informatique INRAE](#)
- [Site web de la plate-forme](#)

Objet et domaine d'application : La présente charte a pour objet de définir les conditions d'utilisation et les règles de bon usage des serveurs, cluster de calcul, espaces disques et services de la plate-forme bio-informatique GenoToul.

Portée de la charte : La présente charte est subordonnée à la [charte informatique](#) d'INRAE disponible sur le site web d'INRAE, ainsi qu'à la [politique de confidentialité des données](#) de la plate-forme GenoToul-Bioinfo.

Mail du support : support.bioinfo.genotoul@inrae.fr

B) Objectifs

La plate-forme bio-informatique GenoToul fournit un accès à :

- Des ressources incluant un cluster de calcul et le stockage associé, des banques de données et des logiciels « open source ».
- Des services incluant l'organisation de cycles d'apprentissages, l'hébergement de sites web et de machines virtuelles, le monitoring et les statistiques associés aux ressources et aux services, une expertise et un support scientifique à des projets relevant de la biologie ou de la bio-informatique.

L'infrastructure mise à disposition par la plate-forme est destinée aux biologistes, bioinformaticiens et statisticiens (en priorité de la région Occitanie et INRAE quel que soit la région), ayant des besoins de calcul intensif de volumétrie moyenne et d'un environnement bio-informatique (logiciels, banques de données, support...).

La plate-forme bio-informatique GenoToul dispose d'une infrastructure dimensionnée pour accueillir un nombre important, mais limité, d'utilisateurs. Le nombre maximum de comptes ouverts (appelé seuil dans la suite du document) est défini par l'équipe de direction de la plate-forme en fonction des moyens matériels et humains disponibles.

Pour maintenir la qualité du service rendu la plate-forme se réserve le droit :

- de renvoyer certains projets sur les centres de calcul régionaux, nationaux ou autres plateformes offrant les mêmes services
- de suspendre la création de nouveaux comptes.

La plate-forme bioinfo GenoToul n'a pas la **certification l'autorisant à héberger des données de santé** ou à caractères personnels. Chaque utilisateur est soumis à la loi de protection des données personnelles lorsqu'il stocke et/ou analyse ses données sur les serveurs de la plate-forme GenoToul Bioinfo (par exemple des données de séquençage humain).

C) Conditions d'accès à la plate-forme

1) Accès à l'infra + demandes exceptionnelles

Pour bénéficier des ressources informatiques de l'infrastructure GenoToul, un utilisateur doit pouvoir démontrer que son projet est dans le domaine de la bio-informatique ou des sciences du vivant.

Pour demander la création d'un compte, un utilisateur doit compléter le [formulaire de création de compte](#) accessible via le site web de la plate-forme. Ce formulaire est accessible tant que les ressources humaines et matérielles le permettent.

L'équipe de direction de la plate-forme se réserve le droit de créer des comptes pour des personnes en collaboration avec des membres de la plate-forme même en cas de suspension d'ouverture des comptes.

L'attribution d'un compte est soumise à :

- une validation par l'équipe de la plate-forme ;
- l'acceptation de la présente charte, de la charte informatique d'INRAE et de la politique de confidentialité des données de la plate-forme GenoToul-Bioinfo ;
- la possession d'un mail institutionnel (académique ou privé) français. Seuls les utilisateurs collaborant avec les membres de la plateforme sont autorisés à avoir un compte s'ils travaillent à l'étranger.

Le compte est strictement personnel, incessible et temporaire : voir les conditions de prolongation et de fermeture d'un compte (ci-dessous paragraphe [C.4](#)).

2) Services liés au compte utilisateur

Chaque compte utilisateur sur la plate-forme donne droit à :

- **l'accès au cluster de calcul via les serveurs de login** permettant :
 - **l'utilisation de banques de données** du domaine : certaines sont systématiquement indexées pour faciliter l'utilisation d'outils en bio-informatique. Elles sont mises à jour automatiquement et régulièrement et seule la version n-1 est conservée ;
 - **l'utilisation des logiciels** du domaine: ils sont mis à disposition d'un

- environnement de travail en bio-informatique ;
- **Un accès aux espaces disques suivants :**

`/home` : espace disponible pour stocker vos fichiers de configuration (volumétrie et conditions d'utilisation: cf [Default resources FAQ : Disk space available](#)).

`/work/user` : espace de travail temporaire. Tous les nœuds du cluster peuvent lire et écrire sur cet espace disque. Les fichiers localisés sur cet espace peuvent être effacés automatiquement si les administrateurs système de la plate-forme le jugent nécessaire (volumétrie et conditions d'utilisation: cf [Default resources FAQ : Disk space available](#)).

`/save/user` : espace disponible pour stocker les données à conserver. Les nœuds du cluster ne peuvent y accéder qu'en lecture. Cette espace disque est automatiquement répliqué sur un autre système de stockage (volumétrie et conditions d'utilisation: cf [Default resources FAQ : Disk space available](#)).

- un quota d'heures de calcul pour les utilisateurs académiques et pour les utilisateurs d'entreprises privées. Au-delà de ces quotas, les utilisateurs devront faire une demande exceptionnelle. Les utilisateurs du secteur privé seront alors facturés au tarif indiqué sur le site de la plate-forme (volumétrie et conditions d'utilisation : cf [Default resources FAQ: Default CPU time](#)).
- **l'accès aux services web de la plate-forme GenoToul-Bioinfo comprenant:**
 - L'accès à l'instance [Galaxy](#) ;
 - L'accès aux [services de monitoring](#) du cluster de calcul ;
 - L'accès au [self-service](#) de la plate-forme ;
 - L'accès à la [liste des banques de données](#) disponibles
- **en cas de besoin** les utilisateurs peuvent faire des **demandes exceptionnelles** via le [formulaire de demande exceptionnelles](#) sur notre site web. Celle-ci sera examinée par le personnel de la plate-forme, et pourra, selon les cas, impliquer l'établissement d'un devis.
Ces demandes peuvent viser à :
 - augmenter l'espace de stockage ;
 - disposer de machines virtuelles ;
 - héberger des sites web (et leurs sites de statistiques si demandé) ;
 - augmenter le quota de calcul ;
 - accéder temporairement à des comptes de formations
 - Créer un espace disque de type « projet » partagé entre un groupe d'utilisateurs.

3) Engagements de la plate-forme

La plate-forme s'engage :

- **à assurer un support** : dans les limites de la disponibilité des agents. L'accès au support se fait via le [formulaire support](#) accessible par le site web ;
- **à informer les utilisateurs** : notamment à prévenir aussi tôt que possible les utilisateurs des maintenances matérielles et systèmes planifiées impliquant une coupure de service.

L'information aux utilisateurs sur l'administration des services (changements de politique, mise en place de nouveaux services, interruptions de service...) se fait au moyen de la newsletter et de mails d'avertissements (à l'adresse mail communiquée par l'utilisateur lors de sa demande de création de compte).

La plate-forme s'engage à faire de son mieux pour limiter au maximum les interruptions de service, mais ne peut pas garantir la disponibilité permanente de ses services. Le personnel de la plate-forme ne peut en aucune façon être considéré comme responsable ni de la perte des données d'un utilisateur, ni de leur diffusion.

4) Durée de validité d'un compte/conditions de fermeture

a) Durée de validité par défaut d'un compte

- Un compte pour les personnels temporaires a une durée de validité par défaut correspondant à la durée de leur contrat déclarée lors de la demande de création de compte. Il sera soumis aux conditions de fermeture d'un compte (voir-ci-dessous [C.4.c](#)).
- Un compte pour les personnels permanents ou titulaires a une durée de validité soumise aux conditions de fermeture d'un compte (voir-ci-dessous [C.4.c](#)).

b) Durée de validité d'un mot de passe

Le mot de passe du compte expire tous les 12 mois et doit être renouvelé via le [self-service](#).

c) Fermeture d'un compte

Un compte sera fermé dans les cas suivants :

- date de fin de contrat déclarée lors de la création du compte atteinte. Attention dans le cas des comptes « privés » (i.e. non académique), la durée du compte est fixée à un mois par défaut (avec un quota de 500h CPU). Elle pourra être prolongée en achetant des heures CPU via notre formulaire de demandes exceptionnelles : [Resources request - genotoul-bioinfo](#) ;
- non-utilisation d'aucun des services de la plate-forme pendant 1 an (pas de connexion au cluster de calcul, aux machines virtuelles, à l'administration des sites web hébergés ou à Galaxy, mot de passe expiré depuis au moins un an) ;
- mail institutionnel non valide ;
- si l'utilisateur n'a pas respecté la présente charte ou la charte informatique d'INRAE ;
- à la demande de l'utilisateur ;
- non respect des contrats établis avec la plate-forme.

Tout utilisateur sera averti de l'expiration de son compte et/ou de son mot de passe 30 jours et 15 jours avant la date d'expiration à l'adresse e-mail fournie lors de la création de son compte. Les données de l'espace de stockage "save" et leur réplication seront effacées 3 mois après la fermeture du compte. Les données temporaires de l'espace de stockage "work" seront supprimées dès la fermeture du compte.

Attention : Les utilisateurs dont le compte va être fermé sont responsables du transfert des données avant la date limite sur un autre compte, un espace projet ou un autre type de stockage en collaboration avec leur encadrant ou collaborateur.

d) Prolongation de compte

Un compte peut être prolongé au-delà de sa date de fin prévue dans les cas suivants :

- prolongation du contrat de l'utilisateur sans changement de situation ;
- en cas de changement de situation, si le seuil du nombre d'utilisateurs n'est pas atteint et si les informations de comptes sont mises à jour via le formulaire de création et de mise à jour de compte ;
- en fin de contrat une extension temporaire est possible (maximum 2 mois) tant que le mail institutionnel reste valide ;

- si le compte a accès à des services loués sur la plate-forme (stockage, CPU, machines virtuelles, la partie administration des sites web hébergés).

5) Supervision du fonctionnement

Les connexions et tentatives de connexion sur le cluster de calcul sont enregistrées. Dans certains cas, les commandes passées par les utilisateurs et/ou des fichiers appartenant à des utilisateurs peuvent être stockées à des fins d'analyse ultérieure dans le cadre de la loi.

Les administrateurs système peuvent :

- interrompre sans préavis tout traitement de données lancé directement sur les serveurs d'accès ;
- interrompre tout processus tournant sur les serveurs d'accès ;
- interrompre toute tâche utilisateur dans le cas où une utilisation excessive des ressources nuit au bon fonctionnement du cluster de calcul ;
- mettre fin aux sessions de travail trop longtemps inactives ;
- effacer les données sur les espaces temporaires (/scratch, /tmp, /work ...) en cas de nécessité.

D) Devoirs de l'utilisateur

1) Sobriété numérique

L'utilisateur a le devoir de limiter au maximum l'utilisation des CPU, mémoire et du stockage pour limiter l'empreinte carbone de ses analyses. Nous mettrons régulièrement sur cette page quelques conseils dans cet objectif : <https://bioinfo.genotoul.fr/index.php/resources-2/eco-friendly-practices/>

2) Soumission de jobs de calcul

Le cluster de calcul dispose d'un gestionnaire de tâches. Les utilisateurs doivent soumettre leurs traitements en utilisant le gestionnaire de tâches. Le contournement du gestionnaire de tâches peut donner lieu à une suspension immédiate du programme soumis et à une éventuelle suspension voire suppression de l'accès aux ressources du cluster de calcul.

L'utilisateur est responsable des tâches qu'il lance tant sur le serveur de login que sur le cluster de calcul. Il doit les suivre et les supprimer en cas de problème.

Les serveurs d'accès sont réservés exclusivement à la connexion, au transfert de données, à la compilation, au test rapide de lignes de commandes et à la soumission de jobs sur le cluster de calcul.

Tout traitement de données lancé directement sur les serveurs d'accès sera systématiquement interrompu sans préavis ni information par les administrateurs système.

3) Stockage des données/sauvegarde

Chaque utilisateur est tenu de n'utiliser que la quantité d'espace disque qui lui est strictement nécessaire, et d'utiliser efficacement les moyens de compression et d'archivage.

Les fichiers temporaires de calcul doivent être supprimés à l'issue du calcul.

Il est de la responsabilité de l'utilisateur de gérer ses données (organisation, volumétrie, pertinence, ancienneté).

Vous pourrez trouver le plan de gestion de données de la plateforme GenoToul Bioinfo sur ce lien :

<https://entrepot.recherche.data.gouv.fr/dataset.xhtml?persistentId=doi:10.57745/ATSYED>

4) Utilisation des logiciels/banques

L'installation et l'utilisation de logiciels ou de banques par les utilisateurs doit se faire dans le respect des dispositions du code de la propriété intellectuelle. Les utilisateurs sont tenus de vérifier les licences correspondantes. En particulier, l'utilisateur doit être en mesure de présenter une licence en règle en cas de demande. Cela vaut également pour les utilisateurs de Galaxy.

L'installation de logiciels ou utilitaires pouvant porter atteinte à l'intégrité des systèmes n'est pas autorisée.

L'utilisateur s'engage à ne pas télécharger les logiciels installés sur l'infrastructure.

Les demandes d'installation multi-utilisateurs de logiciels ou de banques se font en utilisant les pages du site web <http://bioinfo.genotoul.fr> suivantes :

- [Bank installation](#).
- [Software installation](#).

L'installation sera réalisée par l'équipe d'administration système de la plate-forme dans la limite des possibilités : pertinence, temps, ressources nécessaires, dépendances logicielles.

5) Informations personnelles et mot de passe

Chaque utilisateur s'engage à communiquer au support (support.bioinfo.genotoul@inrae.fr) tout changement ou complément d'informations le concernant (mail, laboratoire, numéro de téléphone, groupe...). L'utilisateur peut vérifier ses informations de compte sur le [self-service](#) de la plate-forme.

Les informations personnelles fournies par l'utilisateur lors de la création de son compte sont soumises à la [politique de confidentialité des données](#) de la plate-forme GenoToul-Bioinfo .

Chaque utilisateur s'engage à protéger son compte par un mot de passe, qu'il renouvellera tous les ans.

Tout compte utilisateur est incessible (l'utilisateur s'engage à ne pas communiquer son mot de passe). L'utilisateur est responsable des dommages qui pourraient provenir d'une mauvaise utilisation de son compte s'il est établi qu'il en a donné le mot de passe.

De plus s'il est établi qu'un utilisateur a prêté son compte, celui-ci pourra être fermé par la plate-forme.

6) Respect des directives

Les utilisateurs s'engagent à se conformer aux différentes directives communiquées par la plate-forme grâce aux moyens cités dans le paragraphe [C.3](#) (engagement de la plate-forme). Ils doivent prévenir les administrateurs de la plate-forme de toute anomalie détectée, tant sur le cluster de calcul que sur une machine distante servant à travailler sur le cluster de calcul (via le [formulaire de support](#))

7) Respect du caractère confidentiel

Lors de leur création, les répertoires utilisateurs ont les droits suivants :

- home: drwxr-x--x lecture, écriture et exécution pour le user, lecture et exécution pour le groupe, exécution pour tous (permet l'accès page web au public_html) ;
- save: drwxr-x--- comme le home mais aucun droit pour tous ;
- work: drwxr-x--- comme le home mais aucun droit pour tous.

Il appartient à l'utilisateur de gérer ses droits comme il le souhaite en changeant les droits par défaut (commande `chmod`, `acl`). Pour partager des données, vous pourrez utiliser le mode opératoire dans notre FAQ (Linux / « How to change permission on file or folder »)

Les fichiers sont privés. Toute tentative de lecture ou de copie de fichiers d'un autre utilisateur sans son autorisation est répréhensible. De même que l'interception de communications entre utilisateurs.

Dans le cas de travaux confidentiels, il est de la responsabilité de l'utilisateur de garantir leur confidentialité.

8) Données à caractère personnel

L'utilisateur d'un compte sur la plate-forme GenoToul-Bioinfo déclare être informé qu'en tant que « responsable de traitement » au sens de la loi relative au Règlement Général sur la Protection des Données (RGPD), il lui incombe de respecter et/ou d'effectuer auprès de la CNIL les formalités préalables à la mise en œuvre de tout traitement comportant des données à caractère personnel.

9) Hébergement de machines virtuelles

Les utilisateurs d'un service hébergé doivent également respecter la Charte de la plateforme Genotoul Bioinfo.

L'administration système et la sécurisation de cette (ces) machine(s) incombent à l'hébergé. Dans ce cas, les rôles des deux parties sont les suivants :

- Pour la plateforme : configuration des interfaces réseau, mise à jour de l'infrastructure matérielle, installation des logiciels de monitoring et de sécurité.
- Pour l'hébergé : administration, sécurisation (mises à jours de sécurité), respect du RGPD. Les outils installés par la plateforme ne doivent pas être désinstallés.

E) Propriété et valorisation des résultats

Le laboratoire de recherche a la maîtrise de la valorisation des résultats, toutefois, les équipes de recherche s'engagent à remercier la plate-forme bio-informatique de Toulouse Occitanie dans leur(s) publication(s) et communication(s), lorsqu'elle(s) traite(nt) d'un sujet ayant bénéficié des ressources de la plate-forme, en utilisant la phrase suivante : "We are grateful to the GenoToul bioinformatics platform Toulouse Occitanie (GenoToul-Bioinfo, doi: 10.15454/1.5572369328961167E12) for providing help and/or computing and/or storage resources".

Even non-french speaking users have to accept this charter and the INRA IT charter.