

Rapport de Stage

CENTRE NATIONAL DE LA RECHERCHE SCIENTIFIQUE



cnrs

PIERRE BURNIER

ÉTUDIANT BTS SIO 2A - MARIE CURIE, MARSEILLE 13005

STAGE EFFECTUÉ DU 6 JANVIER AU 13 FEVRIER 2026

SOMMAIRE

Introduction et remerciements	1
Présentation de l'organisme et de son équipe	2
Contexte et Objectifs du stage	3
Les missions effectuées pendant le stage	4
Analyse des résultats et difficultés	5
Bilan personnel et professionnel	6
Conclusion	7
Annexes	8

1. Remerciements

Avant d'entamer le récit de cette expérience, je tiens à remercier toutes les personnes qui ont rendu ce stage possible et enrichissant au sein du Lycée Marie Curie.

Mes remerciements s'adressent tout d'abord à M. Paul Bruno, mon professeur principal, pour m'avoir orienté vers cette opportunité d'excellence au sein du CNRS.

Je remercie chaleureusement Mme Florence Bansept, responsable de l'équipe M3G, pour m'avoir accueilli au sein de son groupe de recherche et m'avoir permis d'évoluer dans un environnement aussi stimulant.

Un grand merci à mon tuteur, M. Léon Espinosa. Son expertise sur l'infrastructure du serveur de calcul et ses conseils sur les besoins des utilisateurs ont été essentiels à la réussite de mon projet de développement.

Je salue également toute l'équipe M3G : Alessandra, Florian, Vitor, Matthew, Julianna et Julien. Un merci particulier à mon binôme, Abdelli Benazzouz, pour son travail sur la documentation technique et SSH, essentiel à la pérennité de l'outil.

Enfin, je remercie M. Matthieu Bergé pour nous avoir permis de tester l'outil sur ses données réelles, validant ainsi l'utilité concrète de notre travail.

Introduction

Dans le cadre de ma formation en BTS SIO (Services Informatiques aux Organisations) au Lycée Marie Curie, j'ai choisi d'effectuer mon stage de six semaines au sein du CNRS de Marseille. Cette immersion au Laboratoire de Chimie Bactérienne (LCB) a représenté un défi de taille : adapter les standards du développement logiciel aux exigences de la recherche fondamentale. Mon objectif n'était pas seulement de coder, mais de réaliser une véritable transformation numérique pour des utilisateurs dont l'informatique n'est pas le cœur de métier.

L'enjeu central de ma mission a été de concevoir une passerelle entre la puissance de calcul brute de l'institut et le quotidien des chercheurs. En adoptant une démarche d'ingénierie logicielle, j'ai dû traduire des besoins scientifiques complexes en une interface ergonomique et fonctionnelle. Ce rapport retrace l'intégralité de mon processus, depuis l'analyse des besoins de l'équipe M3G jusqu'au déploiement final de la solution, tout en mettant en lumière l'importance cruciale de la documentation pour assurer la maintenance de l'outil.

2. Présentation de l'organisme

Le CNRS structure la recherche française à travers des pôles d'excellence comme l'Institut de Microbiologie de la Méditerranée (IMM). Au cœur de cet institut, le Laboratoire de Chimie Bactérienne (LCB) se dédie à l'étude des systèmes bactériens. J'ai été intégré à l'équipe M3G (Mathematical Modeling of Microbiomes), dirigée par Florence Bansept. La force de cette équipe réside dans sa pluridisciplinarité, où la biologie de terrain rencontre la modélisation mathématique pour décrypter la complexité des écosystèmes microbiens. On y retrouve des thématiques variées : Florian Labourel travaille sur l'écologie évolutive, tandis que Vitor Marquioni Monteiro et Julien Vogelweid se concentrent sur la modélisation du microbiote intestinal. Mon tuteur, Léon Espinosa, assure le lien technique en pilotant notamment la segmentation d'images et les simulations numériques.

Cependant, les besoins en ressources de calcul ne s'arrêtent pas aux frontières de l'équipe M3G. L'IMM dispose d'un serveur de calcul haute performance (HPC), équipé de 18 cartes graphiques NVIDIA A40. Cette puissance attire des chercheurs de multiples unités de recherche du campus qui ont besoin de traiter des volumes de données expérimentales importants. Le problème identifié était universel au sein du laboratoire : l'accès à ce monstre de puissance se faisait exclusivement via terminal SSH, une barrière technique extrêmement chronophage pour la majorité des biologistes.

Mon projet a donc pris une dimension transversale : concevoir une interface ergonomique capable de "démocratiser" l'accès au GPU. En simplifiant le lancement de scripts personnalisés, la gestion d'environnements Conda et l'automatisation de la segmentation, l'outil développé a permis de transformer une ressource complexe en un service intuitif, optimisant ainsi le flux de travail de l'ensemble de l'organisme.

Présentation de l'équipe



**FLORENCE
BANSEPT**

Cheffe d'équipe



**LÉON
ESPINOSA**

Ingénieur
Chercheur



**ALESSANDRA
FRIEDRICH
LUTZ**

Postdoc



**FLORIAN
LABOUREL**

Postdoc



**VITOR
MARQUIONI
MONTEIRO**

Postdoc



**MATTHEW
JONES**

Doctorant



**JULIANA
CICCHETO**

Doctorante



**JULIEN
VOGELWEID**

Stagiaire Master 2



**PIERRE
BURNIER**

Stagiaire BTS



**ABDELLI
BENAZZOZ**

Stagiaire BTS

3. Objectifs du Stage

Le contexte de ce stage s'inscrit dans une phase de transition numérique majeure pour l'Institut de Microbiologie de la Méditerranée. Alors que la recherche en biologie génère des volumes de données de plus en plus massifs — notamment via l'imagerie haute résolution et les simulations stochastiques — les méthodes de traitement restaient jusqu'alors artisanales. La pratique courante consistait soit à saturer les ressources locales des postes de travail, soit à solliciter ponctuellement un ingénieur spécialisé pour manipuler le cluster de calcul. Ce mode de fonctionnement créait une rupture dans le flux de recherche : le scientifique perdait la main sur son cycle d'analyse dès que celui-ci dépassait les capacités de son ordinateur personnel.

L'objectif premier de ma mission a donc été de restaurer cette continuité en développant une interface "façade". Il ne s'agissait pas simplement de créer un site web, mais de concevoir une couche d'abstraction logicielle capable de masquer la complexité du système d'exploitation Linux et de la gestion des ressources GPU. L'enjeu fonctionnel était triple : permettre la création dynamique d'environnements de travail isolés via Conda, offrir une plateforme d'exécution pour des scripts Python personnalisés sans risque de conflit système, et automatiser les tâches répétitives de segmentation d'images.

Enfin, l'ambition finale de ce projet est d'assurer une autonomie totale et durable au sein de l'institut. L'interface a été pensée pour devenir un outil de transmission : sa simplicité d'utilisation doit permettre aux chercheurs permanents de former instantanément les nouveaux arrivants, qu'ils soient doctorants ou stagiaires. En abaissant la barrière technique à l'entrée, mon travail visait à transformer le serveur de calcul en un outil pédagogique et collaboratif, garantissant que le savoir-faire technique ne soit plus l'apanage de quelques experts, mais un levier accessible à tous pour la pérennité des recherches du laboratoire.

4. Les missions effectuées pendant le stage

Ma mission principale a consisté à transformer des procédures manuelles complexes en un flux de travail (workflow) automatisé et sécurisé. En tant que développeur, j'ai dû concevoir une logique de génération dynamique de scripts. L'enjeu technique était de traduire les saisies des utilisateurs via une interface PHP/Bootstrap en fichiers de configuration interprétables par le serveur (Slurm, YAML, Bash).

Au lieu de laisser l'utilisateur manipuler directement le terminal, l'application agit comme un "chef d'orchestre" : elle collecte les paramètres (choix du GPU, chemins de données, bibliothèques Python), les injecte dans des templates de code préétablis, et fournit un guide interactif pour l'exécution finale. Cette approche par "formulaire guidé" réduit drastiquement les erreurs de syntaxe et garantit que les ressources massives du serveur (les 18 cartes A40) sont utilisées de manière optimale et conforme aux règles de l'institut.

Enfin, j'ai organisé une phase de déploiement active, passant par une présentation magistrale devant un auditoire de chercheurs et de stagiaires, suivie d'un support technique de proximité. Cette approche hybride, entre ingénierie et accompagnement, a permis de garantir que l'outil ne soit pas seulement une prouesse technique, mais un service partagé, compris et utilisé par l'ensemble des acteurs de l'institut.

1

Développement de l'interface d'abstraction

Conception d'un tableau de bord ergonomique avec Bootstrap et JavaScript pour masquer la complexité du système Linux.

2

Automatisation du pipeline de segmentation

Intégration du script Python de segmentation bactériologique via une interface de saisie générant automatiquement les scripts d'appel.

3

Gestion des environnements virtuels (Conda)

Mise en place d'un générateur de fichiers YAML. L'utilisateur sélectionne ses dépendances via le formulaire, et l'outil produit le fichier de configuration ainsi que les commandes pour créer et activer l'environnement de travail isolé sur le cluster.

4

Orchestration et soumission Slurm

Création d'un moteur de génération de scripts .slurm. En fonction des besoins en puissance (GPU/RAM), l'application calcule et rédige le fichier de soumission, puis guide l'utilisateur dans l'envoi de la tâche via la commande sbatch.

5

Formation et support aux utilisateurs

Présentation de l'outil final devant un groupe de chercheurs et stagiaires de l'IMM. Cette mission s'est poursuivie par un accompagnement individualisé pour l'installation et la prise en main de l'interface, assurant ainsi l'autonomie des utilisateurs et la pérennité de l'outil.

5. Analyse des résultats et difficultés

La double barrière de la communication

Au-delà du code, ce stage a imposé de franchir une double barrière linguistique et conceptuelle. Il a fallu d'une part traduire les besoins métiers de la biologie (segmentation, microbiote) en logique algorithmique, et d'autre part gérer l'internationalisation de l'outil. Le monde de la recherche étant majoritairement anglophone, j'ai implémenté un système de traduction complet pour le site. Ce défi s'est accompagné d'une exigence de "Clean Code" et d'optimisation : mon code devait non seulement être performant, mais aussi lisible et maintenable pour que l'équipe puisse le faire évoluer. L'anticipation des erreurs utilisateurs a été le fil rouge de ma documentation, où chaque "pépin" potentiel a été répertorié pour garantir une utilisation sereine, même en mon absence.

Des défis techniques et environnementaux

Le développement de l'interface a révélé des problématiques d'interopérabilité majeures, notamment la gestion des différences entre les systèmes Windows, Mac et Linux utilisés par les chercheurs. L'un des obstacles les plus instructifs a été la gestion de la structure de fichiers locale. Un navigateur web, par sécurité, ne peut pas interagir directement avec le disque dur de l'utilisateur et place systématiquement les fichiers dans le dossier "Téléchargements". J'ai donc dû concevoir un système de guidage extrêmement précis pour accompagner l'utilisateur dans le déplacement manuel de ses scripts et environnements, afin de respecter l'arborescence rigoureuse exigée par le serveur de calcul. Cette contrainte a nécessité une réflexion poussée sur l'ergonomie pour transformer une limitation technique en un processus utilisateur fluide et sans erreur.

Un bilan opérationnel probant

L'analyse des résultats confirme le succès de la solution : l'outil est aujourd'hui en production et a radicalement transformé le quotidien du laboratoire. Le gain de temps est massif, mais c'est surtout l'autonomie retrouvée des chercheurs qui marque la réussite du projet. La barrière du terminal SSH n'est plus un frein à la science. La preuve la plus gratifiante de ce succès reste son adoption immédiate par des membres extérieurs à l'équipe initiale. Par exemple, Matthieu Bergé a non seulement intégré l'outil à son workflow, mais il forme désormais ses propres stagiaires à son utilisation. Voir mon travail devenir un vecteur de transmission de savoir au sein d'un institut comme le CNRS est la meilleure validation de la pertinence de mon approche.

5. Bilan personnel et professionnel

Ce stage au CNRS a profondément fait évoluer ma vision du monde professionnel et de mon avenir de développeur. L'aspect le plus marquant a été l'immersion dans un environnement d'excellence, entouré de profils d'une grande compétence. J'ai découvert que la recherche n'est pas un monde statique, mais un univers sans routine où s'enchaînent expérimentations, séminaires techniques et présentations orales. Cette dynamique m'a prouvé qu'un métier passion, centré sur la découverte, est un moteur puissant pour l'épanouissement personnel. La "preuve sociale" et le prestige associés à une thèse sont des éléments qui résonnent désormais avec mes propres ambitions de carrière.

Sur le plan de l'employabilité, cette expérience m'a fait prendre conscience de la nécessité de me démarquer. Dans un marché où le niveau Master (Bac+5) devient la norme, s'orienter vers la recherche et le doctorat apparaît comme un levier stratégique pour atteindre un niveau d'expertise supérieur et accéder à des opportunités internationales. L'ambiance collaborative et l'enthousiasme des membres de l'équipe M3G m'ont montré qu'il est possible de concilier exigence scientifique et cadre de travail stimulant. La perspective de travailler sur des sujets de pointe, notamment en Intelligence Artificielle, est une voie que je considère désormais avec beaucoup plus d'intérêt qu'auparavant.

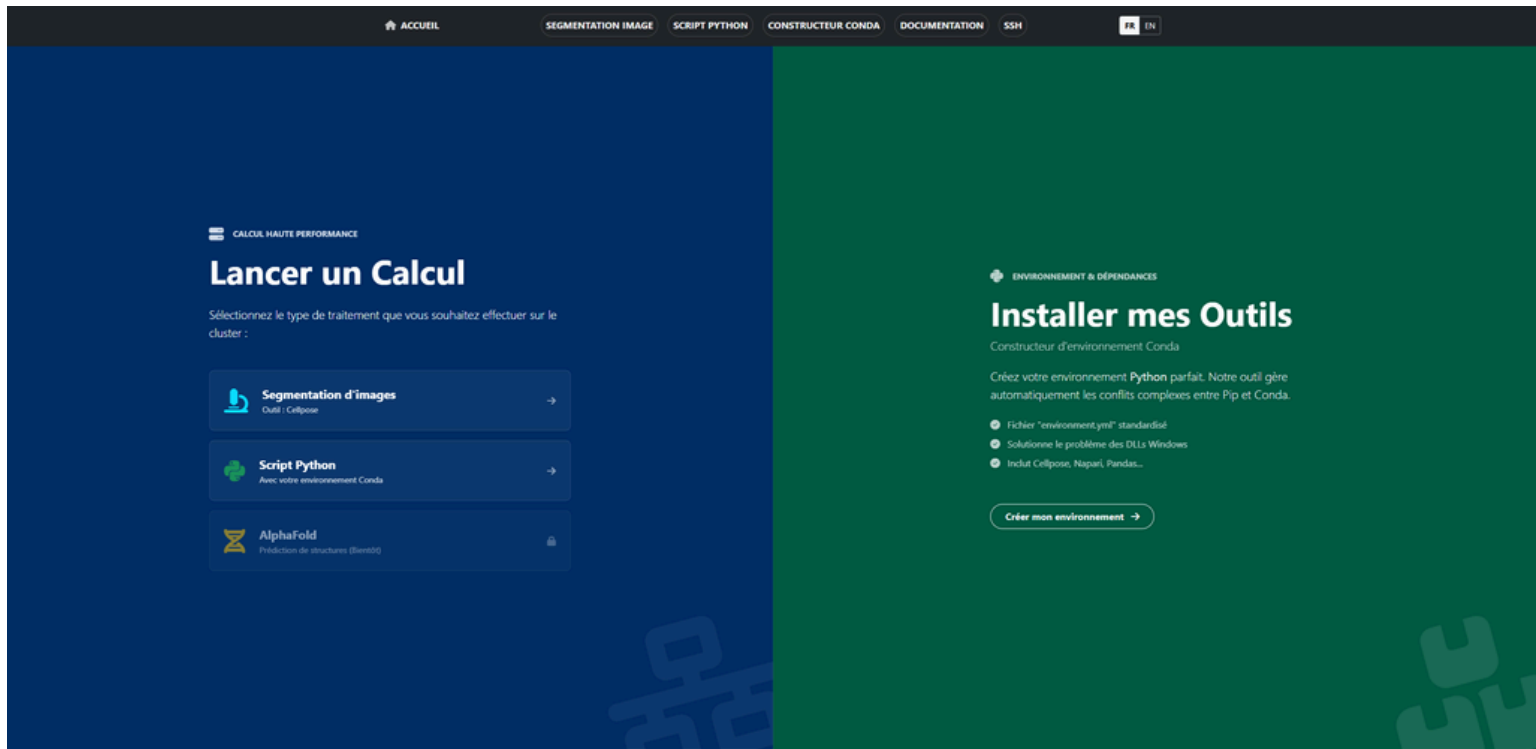
Néanmoins, ce stage m'a aussi apporté une vision lucide sur les réalités du secteur. J'ai observé que le monde de la recherche académique souffre parfois d'un manque de stabilité contractuelle, avec une forte prédominance des contrats post-doctoraux par rapport aux postes permanents en CDI. Cette composante est un facteur que je devrai mettre en balance avec le dynamisme du secteur privé en IA. Quoi qu'il en soit, ces six semaines m'ont permis de confirmer mon attrait pour les défis complexes et les environnements de haute technicité, me laissant repartir avec une certitude : la porte de la recherche est aujourd'hui grande ouverte.

6. Conclusion

Pour conclure, ce stage au sein du Laboratoire de Chimie Bactérienne aura été bien plus qu'une simple application de mes compétences en développement web. Il m'a permis de porter une double casquette : celle de technicien, en concevant un outil d'abstraction pour un cluster HPC de 18 cartes NVIDIA A40, et celle de consultant, en accompagnant des chercheurs vers une plus grande autonomie numérique. Le succès de l'interface auprès des différentes équipes de l'IMM et des collaborateurs témoigne de la pertinence d'une approche orientée utilisateur dans le milieu scientifique.

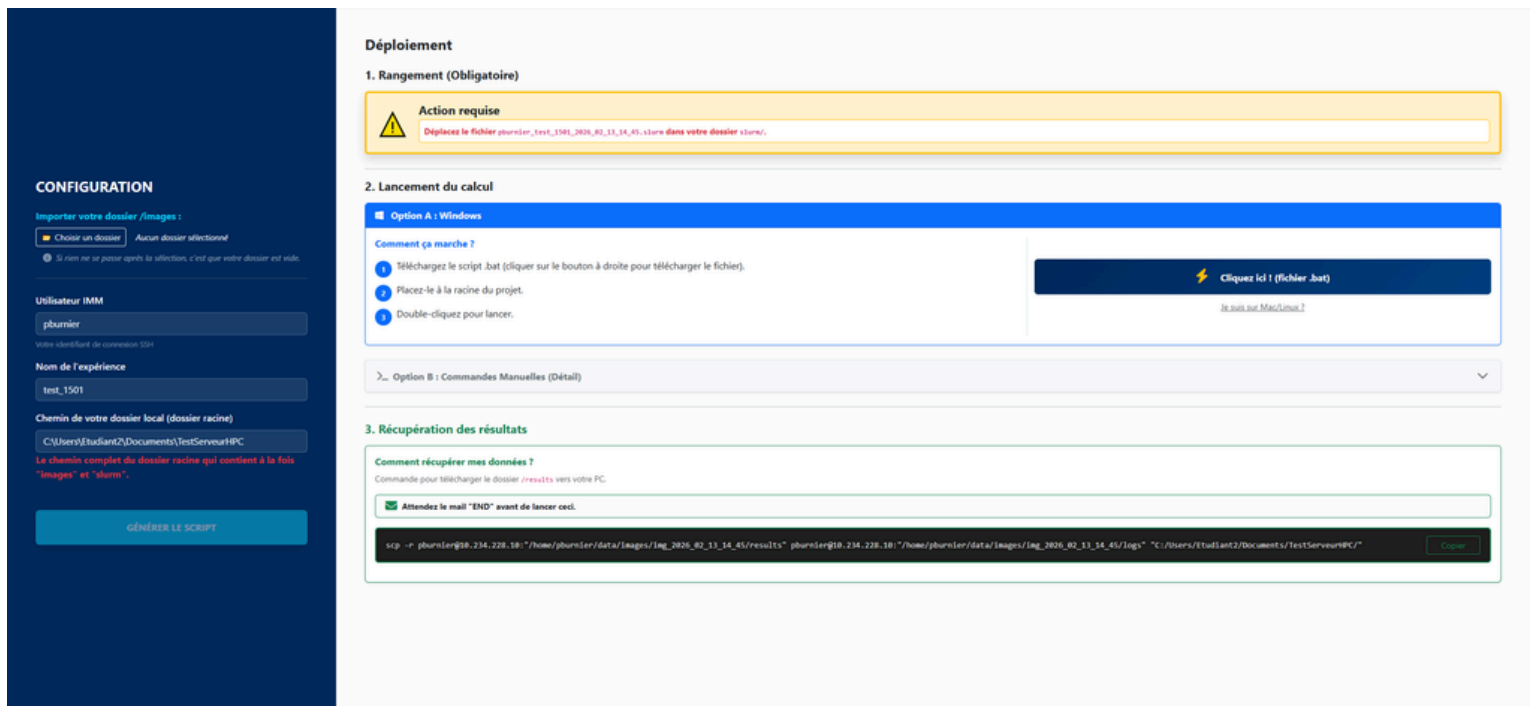
Cette expérience confirme mon envie de poursuivre mes études vers des domaines où l'informatique se met au service de l'innovation et de la compréhension du vivant. Je repars de l'Institut de Microbiologie de la Méditerranée avec des compétences techniques consolidées en architecture client-serveur, mais aussi avec une maturité professionnelle accrue. Que mon futur se dessine dans la recherche académique ou dans l'ingénierie de pointe en IA, les méthodes de travail et la rigueur acquises au CNRS constitueront le socle de mon évolution future.

7. Annexes



Application développée par l'équipe Baisept au Laboratoire de Chimie Bactérienne UMR 7283 AMU-CNRS.

Capture d'écran de la page d'accueil de notre application web



Capture d'écran de la page de segmentation d'image une fois le formulaire rempli

7. Annexes

The screenshot shows the 'GÉNÉRATEUR PYTHON' web application. On the left is a dark sidebar with three sections: 1. IDENTITÉ & PROJET (fields for username 'pburnier', server IP '10.234.228.10', and script name 'test_1501'), 2. CONFIGURATION PYTHON (fields for environment 'base', script name 'hello.py', and local folder 'C:\Users\Etudiant2\Documents\TestServeurIPC'), and 3. ARGUMENTS DU SCRIPT (fields for options 'OPT1=""', 'BATCH_SIZE=32', and 'INPUT=""'). A yellow button 'GÉNÉRER LES SCRIPTS' is at the bottom of the sidebar. The main content area is titled 'Scripts Générés' and contains three steps: 'Étape 1 : Téléchargement' with a download button and a warning box; 'Étape 2 : Lancement Automatique' with a 'Lancer (.bat)' button and instructions; and 'Étape 3 : Récupération des résultats' with a 'Copier' button and a terminal-style output box.

Capture d'écran de la page python de notre application web

The screenshot shows the '1. Paramètres Généraux' section of the application. It features four input fields: 'VOTRE UTILISATEUR IMM' (value: 'ex: pburnier'), 'SERVEUR CIBLE (IP)' (value: '10.234.228.10'), 'NOM DE L'ENVIRONNEMENT' (value: 'conda activate mon_env_imm'), and 'VERSION PYTHON' (dropdown menu showing '3.10 (Stable)'). A green 'Compatibilité Serveur' button is below. The '2. Sélection des Outils' section below has three columns: 'Imagerie & Segmentation' (Cellpose, ND2, TiffFile, Scikit-Image, OpenCV), 'Data Science' (Pandas, Numpy, Matplotlib, JupyterLab, Scikit-Learn, Joblib), and 'Deep Learning' (PyTorch, TensorFlow).

Capture d'écran de la page conda une fois le formulaire rempli

7. Annexes

i Introduction au calcul sur le cluster GPU IMM

Le serveur IMM est un **cluster de calcul HPC** (*High Performance Computing*, ou calcul haute performance). Un cluster HPC est un ensemble de machines puissantes interconnectées, conçues pour exécuter des calculs intensifs en parallèle, bien au-delà des capacités d'un ordinateur personnel. Les traitements sont répartis sur plusieurs ressources matérielles afin d'accélérer les calculs et de permettre l'analyse de volumes de données importants, notamment en imagerie, en intelligence artificielle et en calcul scientifique.

Le cluster IMM dispose de **GPU** et d'unités de calcul spécialisées parfois appelées **HPU** (*Hardware / Habana Processing Unit* selon le contexte). Ces processeurs sont optimisés pour effectuer un très grand nombre d'opérations en parallèle, ce qui les rend particulièrement efficaces pour le **deep learning**, l'analyse d'images et les calculs intensifs.

L'accès au cluster se fait à distance via **SSH** (*Secure Shell*), un protocole sécurisé permettant de se connecter au serveur, de transférer des fichiers et d'exécuter des commandes. Les calculs ne sont pas lancés directement sur la machine : ils sont gérés par **SLURM**, le planificateur de tâches du cluster. SLURM attribue automatiquement les ressources nécessaires (CPU, GPU, mémoire, durée), place les calculs en file d'attente si besoin et garantit un partage équitable du serveur entre les utilisateurs.

Les outils présentés dans cette documentation ont pour objectif de **simplifier l'utilisation du cluster GPU IMM**. Le **générateur Slurm** permet de créer automatiquement des scripts de calcul adaptés à vos besoins, sans avoir à maîtriser toute la syntaxe de SLURM. Le **constructeur Conda** facilite la création d'environnements logiciels reproductibles et compatibles avec le serveur. Ensemble, ces outils vous permettent de vous concentrer sur votre travail scientifique, tout en exploitant efficacement la puissance de calcul du cluster.

☰ Sommaire

1. SEGMENTATION CELLPOSE

- > 1.1 Tutoriel pas à pas
- > 1.2 Comment ça marche ?
- > 1.3 Structure des dossiers
- > 1.4 FAQ & Erreurs

2. SCRIPT PYTHON

- > 2.1 Comment ça marche ?
- > 2.2 Structure des dossiers
- > 2.3 FAQ & Erreurs

3. CONSTRUCTEUR CONDA

- > 3.1 Comment ça marche ?
- > 3.2 Structure des dossiers
- > 3.3 FAQ & Erreurs

i 4. Ressources & Wiki

Capture d'écran de la documentation de notre application web

Windows Mac / Linux

Pré-requis Windows
Utilisez PowerShell en mode Administrateur.
(Clic droit Démarrer > Terminal (Admin))

1 Activer le service SSH (Lancez cette commande uniquement si vous avez ouvert votre terminal en mode administrateur, sinon passez à l'étape 2)
Force le démarrage de l'agent SSH Windows.

```
Set-Service ssh-agent -StartupType Automatic  
Start-Service ssh-agent
```

2 Générer les clés
Appuyez sur **Enter** pour toutes les questions (pas de passphrase).

```
ssh-keygen -t ed25519 -C "votre_email@imm.cnrs.fr"
```

3 Envoyer la clé publique
Remplacez "votre_login" ci-dessous. Entrez le mot de passe quand demandé.

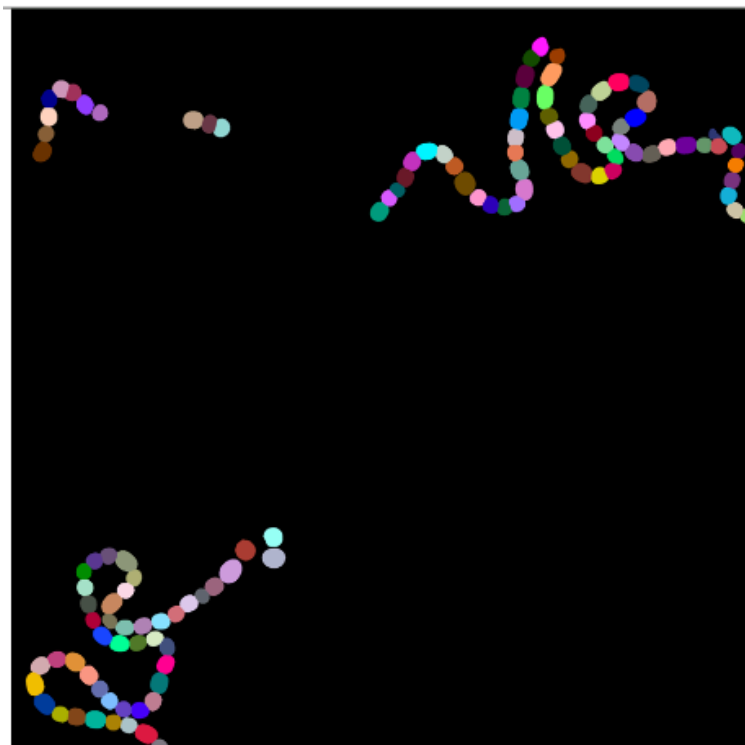
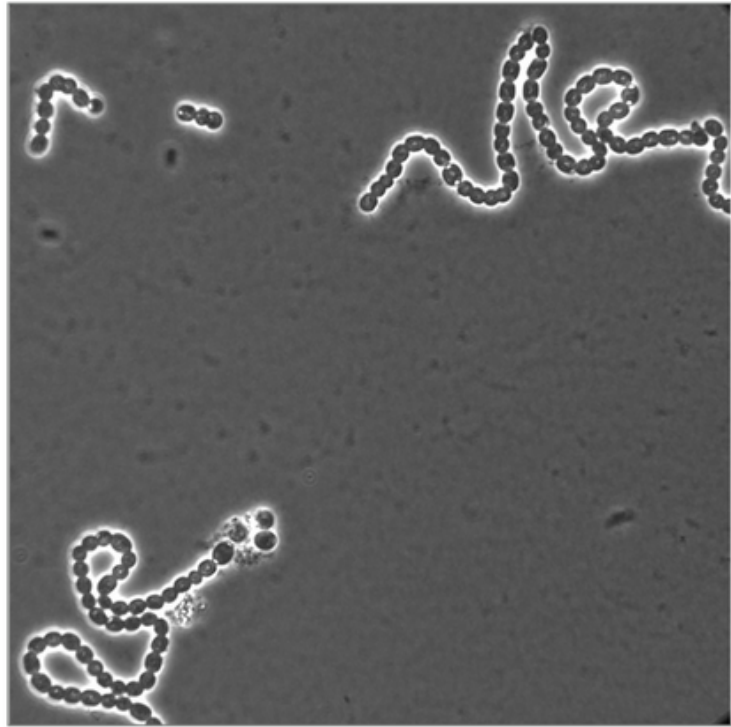
```
type %env:USERPROFILE%\ssh\id_ed25519.pub | ssh votre_login@10.234.228.10 "mkdir  
-p .ssh && cat >> .ssh/authorized_keys"
```

4 Mémoriser la clé
Ajouter la clé à l'agent (session courante)

```
ssh-add %env:USERPROFILE%\ssh\id_ed25519
```

Capture d'écran de la doc ssh

7. Annexes



exemple avant/après segmentation d'image

Logo de l'organisme d'accueil

ATTESTATION DE STAGE

à remettre à la ou au stagiaire à l'issue du stage

ORGANISME D'ACCUEIL

Nom ou dénomination sociale : CNRS
 Adresse : 31, Chemin Joseph AIGUIER, 13009 Marseille

 04 91 16 40 77
 ☎ :

certifie que

LA OU LE STAGIAIRE

Nom : BURNIER Prénom : Pierre
 Né(e) le : 07/03/2005 Sexe : F M
 Adresse : 123.Traverse Parangon Bat GA, 13008 Marseille

 ☎ : 07 81 58 82 55... Mél : pierreburnier07@gmail.com

ÉTUDIANT(E) EN BTS Services informatiques aux organisationsOption SISR SLAM

AU SEIN DE (nom de l'établissement d'enseignement supérieur ou de l'organisme de formation) :
 LYCEE MARIE CURIE

a effectué un stage prévu dans le cadre de ses études

DURÉE DU STAGE

Dates de début et de fin du stage : Du 05/01/2026... au 13/02/2026
 Représentant une **durée totale** de 6 semaines nombre de semaines / de mois
 (rayer la mention inutile).

La durée totale du stage est appréciée en tenant compte de la présence effective de la ou du stagiaire dans l'organisme, sous réserve des droits à congés et autorisations d'absence prévus à l'article L.124-13 du code de l'éducation (art. L.124-18 du code de l'éducation). Chaque période au moins égale à 7 heures de présence consécutives ou non est considérée comme équivalente à un jour de stage et chaque période au moins égale à 22 jours de présence consécutifs ou non est considérée comme équivalente à un mois.

MONTANT DE LA GRATIFICATION VERSÉE À LA OU AU STAGIAIRE

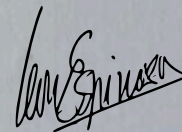
La ou le stagiaire a perçu une gratification de stage pour un **montant total** de euros.

L'attestation de stage est indispensable pour pouvoir, sous réserve du versement d'une cotisation, faire prendre en compte le stage dans les droits à retraite. La législation sur les retraites (loi n°2014-40 du 20 janvier 2014) ouvre aux étudiants dont le stage a été gratifié la possibilité de faire valider celui-ci dans la limite de deux trimestres, sous réserve du versement d'une cotisation. La demande est à faire par l'étudiant(e) dans les deux années suivant la fin du stage et sur présentation obligatoire de l'attestation de stage mentionnant la durée totale du stage et le montant total de la gratification perçue. Les informations précises sur la cotisation à verser et sur la procédure à suivre sont à demander auprès de la Sécurité sociale (code de la Sécurité sociale art. L.351-17 - code de l'éducation art. D.124-9)

Fait à Marseille le 13/02/2026

Nom, fonction et signature de la personne représentant de
 l'organisme d'accueil

Leon Espinosa
 Ingénieur de Recherche, encadrant du stage



¹ Remettre autant d'attestations que d'entreprises fréquentées pour couvrir les semaines de stage réglementaires.

² Pour les personnes candidates se présentant au titre de leur activité professionnelle, cette attestation sera remplacée par des certificats de travail.

Rapport de Stage, CNRS, 06 janvier - 13 février 2026
