

Introduction à Perl pour manipuler simplement ses données

<u>Objectifs</u>	<ul style="list-style-type: none">▶ Acquérir les notions de bases en Perl pour<ul style="list-style-type: none">○ manipuler des données (rassembler, filtrer, reformater)○ automatiser des exécutions de programmes○ générer des rapports▶ A la fin de la formation, les stagiaires seront capables de<ul style="list-style-type: none">○ concevoir des programmes○ adapter des programmes existants○ chercher et récupérer des scripts et des bibliothèques Perl à partir d'Internet
<u>Public concerné</u>	Personnel scientifique, ingénieurs, techniciens doctorants, post-doctorants ayant à traiter/formater de grandes quantités de données (biologistes, chimistes), ASR (fichiers log)
<u>Contenu</u>	<p>Perl est un langage de programmation qui permet de développer rapidement des scripts simples de quelques lignes ou des applications complètes gérant des données complexes.</p> <p>Il est particulièrement adapté pour le traitement de fichiers texte, la génération de rapport et le formatage de données. Il permet de réaliser des contrôles de qualité, la mise en forme des données pour une utilisation ultérieure par des programmes de calculs spécifiques (par exemple statistique avec R). Il est donc utile pour toute personne ayant à traiter des données issues d'instruments scientifiques. Une grande communauté de développeurs ont développé et maintiennent des packages par exemple pour la connexion à des bases de données, la génération de page web dynamiques, la génération de graphiques, la manipulation de fichiers Excel, ou des packages spécifiques à une discipline par exemple bioperl. Cette formation abordera à titre d'exemple l'utilisation de packages pour la manipulation de fichiers excel et la génération de fichiers graphiques.</p> <p>Chaque session de cours est suivi par une application pratique sur PC à partir d'exemples concrets, sur des données si possible du même type que celles auxquelles les stagiaires sont confrontés régulièrement.</p> <p>Les thèmes abordés seront :</p> <ul style="list-style-type: none">▶ la manipulation de données<ul style="list-style-type: none">○ Rassembler des données de plusieurs fichiers, provenant de plusieurs répertoires ou de plusieurs sites○ Extraire des parties spécifiques de ces données○ Réaliser des traitements sur ces données (par exemple des contrôles de qualité)○ Reformater ces données pour les rendre exploitables par d'autres outils▶ La manipulation de programmes<ul style="list-style-type: none">○ Exécuter un ou plusieurs programmes en série (pipeline)○ Garder les traces d'exécution des programmes et les contrôler▶ La génération de rapports<ul style="list-style-type: none">○ Textuels, graphiques ou dans des formats spécifiques tel qu'excel
<u>Minimum requis</u>	Savoir utiliser un ordinateur et un éditeur de texte (par exemple notepad). Pas de connaissance en programmation nécessaire.
<u>Effectif</u>	10 personnes maximum par session
<u>Durée</u>	3 jours
<u>Contacts</u>	Sophie Gallina (sophie.gallina@univ-lille1.fr) et Stefan Gaget (stefan.gaget@good.ibl.fr)