

Tools

search tools

YOUR DATA
[Upload Data](#)
[Download Data](#)

FILES MANIPULATION
[Text Manipulation \(e-learning\)](#)
[Filter and Sort](#)
[Join, Subtract and Group](#)
[Convert Formats](#)
[BED Tools](#)
[Graph/Display Data](#)

SEQUENCES MANIPULATION
[FASTA manipulation](#)
[FASTQ manipulation \(e-learning\)](#)
[SAM/BAM manipulation : Picard \(beta\)](#)
[SAM/BAM manipulation: SAMtools \(e-learning\)](#)
[Fetch Sequences](#)
[Sequences Queries](#)
[VCF Tools](#)

SGS MAPPING
[BWA - Bowtie \(e-learning\)](#)

SNP / INDEL
[NGS: GATK Tools \(beta\)](#)
[Indel Analysis](#)

WELCOME TO GALAXY WORKBENCH



Galaxy is a workbench available for biologists from Sigenae Platform. Galaxy objectives are:

- Make bioinfo Linux tools accessible to biologists.
- Hide the complexity of the infrastructure.
- Allow creation, execution and sharing of workflows.

NEW **BACKUPED DATASET from Sigenae instance of Galaxy before migration**

Access to your old version of Galaxy : [Please clic here](#)

Warnings :

- When you access or reload to your Galaxy webpage, please find all your histories saved in the following menu : "User" / "Saved histories".
- Your data are stored in work/ directory. Consequently, BioInfo Genotoul platform reserves the right to purge all files not accessed since 120 days on work/ disk space.

Sigenae support : sigenae-support@listes.inra.fr

If you have some question about Galaxy, please consult your [FAQ](#)

! How to cite Galaxy workbench ?

History

Copy of 'star- sampl2' shared by 'nazzouzi@toulouse.inra.fr' (active items only)
 187.7 GB

78: report download 7 lines
 format: txt, database: ?
 cp /galaxydata/database/files/015/dataset_15723.dat //work/smaman/GALAXYtoto55//toto; directory : /work/smaman/GALAXYtoto55/ file(s) saved : tata cp /galaxydata/database/files/016/dataset_16565.dat /work/smaman/GALAXYtoto55/tata ;

```
total 8
drwx--x--x 38 smaman BIOINFO 4.8K Jan 13 1
-rw-r--r-- 1 smaman BIOINFO 338 Jan 13 1
lrwxrwxrwx 1 smaman BIOINFO 48 Jan 13 1
-rw-r--r-- 1 smaman BIOINFO 211 Jan 13 1
drwx--x--x 2 smaman BIOINFO 4.8K Jan 13 1
```

72: report download 5 lines
 format: txt, database: ?
 cp /galaxydata/database/files/015/dataset_15723.dat //work/smaman/GALAXYtoto3//toto; directory : /work/smaman/GALAXYtoto3/ file(s) saved : titi

GT GALAXY France

&

Les évolutions de l'instance Sigena



GT GALAXY France



Une « Galaxy » parmi tant d'autres



Galaxy: a comprehensive approach for supporting accessible, reproducible, and transparent computational research in the life sciences

Jeremy Goecks¹, Anton Nekrutenko^{2,3}, James Taylor^{1,3} and The Galaxy Team

Communauté internationale:

- Listes de diffusion
- Wiki
- Twitter
- «Galaxy tour de France »
- Ajouts : Hollande, GB.



Communauté nationale:

- Groupe de travail soutenu par l'IFB
- Fédérer l'activité française autour de Galaxy
- Regrouper les plateformes bioinformatiques de l'IFB.

Membres du GT Galaxy France



- ABiMS : CNRS Station biologique de Roscoff
- MIGALE : INRA Centre de Jouy-en-Josas
- URGI : INRA Versailles
- Institut Curie Paris
- CIRAD SouthGreen
- Sigenae / BioInfo Genotoul : INRA Auzeville

ABiMS Roscoff	Initiation, NGS Cleaning, RNASeq Differential Expression	http://galaxy.sb-roscoff.fr/	Christophe Caron - Alexandre Cormier - Gildas Lecorguille - Pierre Pericard
Institut Curie	ChIP-Seq Analysis	http://nebula.curie.fr/	Alban Lermine
Genotoul / Sigenae	Initiation to Galaxy, SNP calling, RNASeq, sRNASeq	http://galaxy-workbench.toulouse.inra.fr/	Sarah Maman
INRA URGI	Differential expression analysis, Variant detection	http://urgi.versailles.inra.fr/galaxy2	Olivier Inizan
INRA MIGALE	Initiation to Galaxy, NGS Galaxy	http://migale.jouy.inra.fr/galaxy/	Sandra Derozier - Franck Samson
Southgreen	Generalist platform, and crop breeding	gohelle.cirad.fr/galaxy/root/	Jean-Francois Dufayard
INRA PFEM / MetaboHUB	Metabolomics data analysis	https://pfem-galaxy/	Franck Giacomoni

+ INRA Clermont-Ferrand-Theix :
Plateforme technologique membre de l'infrastructure nationale MetaboHub



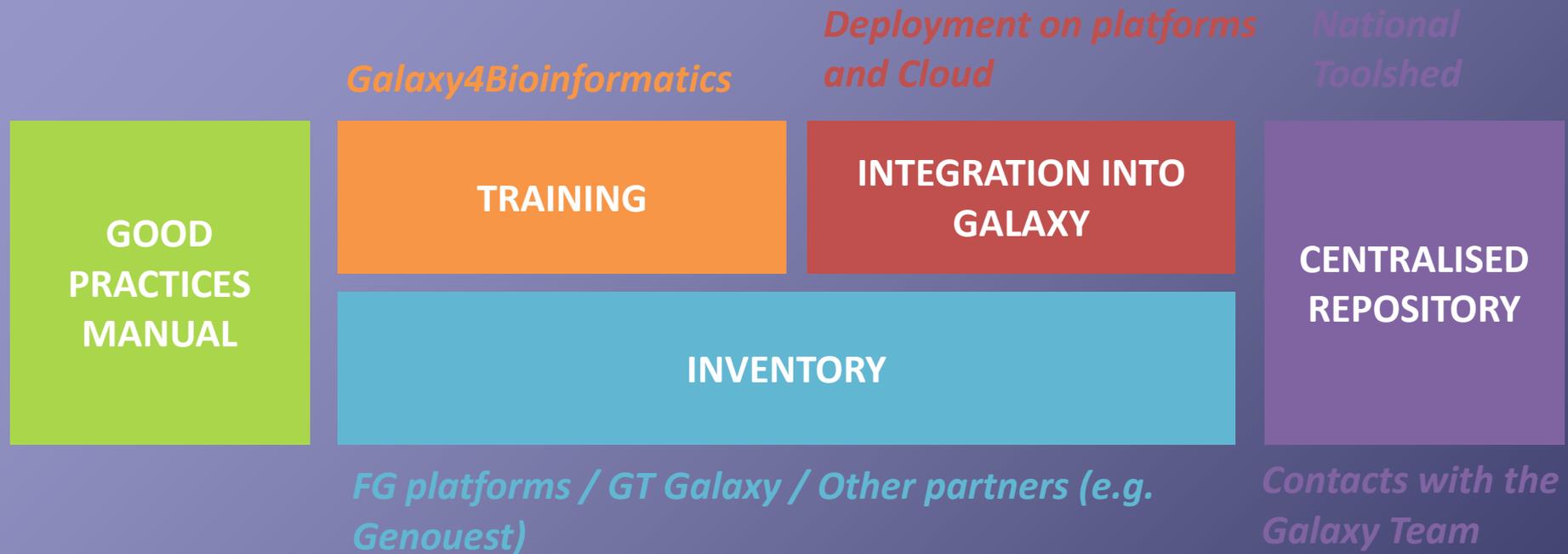
Actions et objectifs du GT Galaxy

Actions de réflexion, d'animation, et de partage:

- Architecture et optimisation
- Documentation collaborative (wiki et liste de diffusion)
- Développement et intégration des outils
- Formations
 - Pour les bio-informaticiens
 - Pour les biologistes
 - Mise en commun des agendas PF

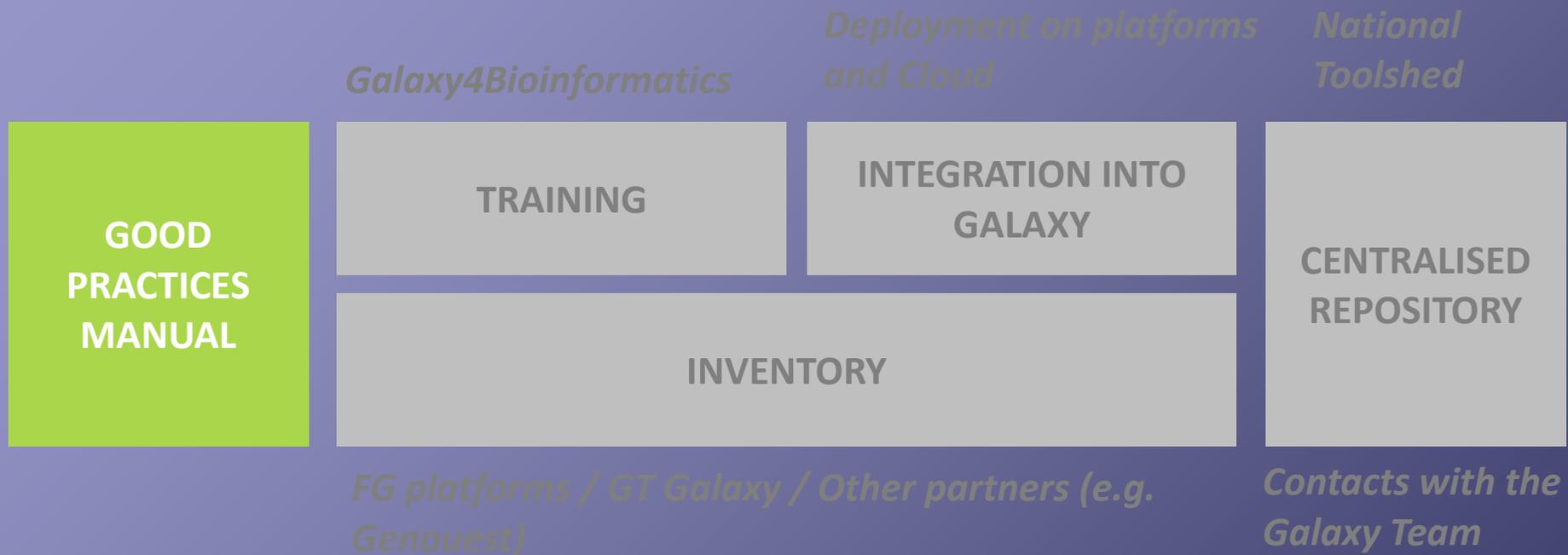
Objectif : Mutualisation des pratiques.

Principales activités du GT Galaxy



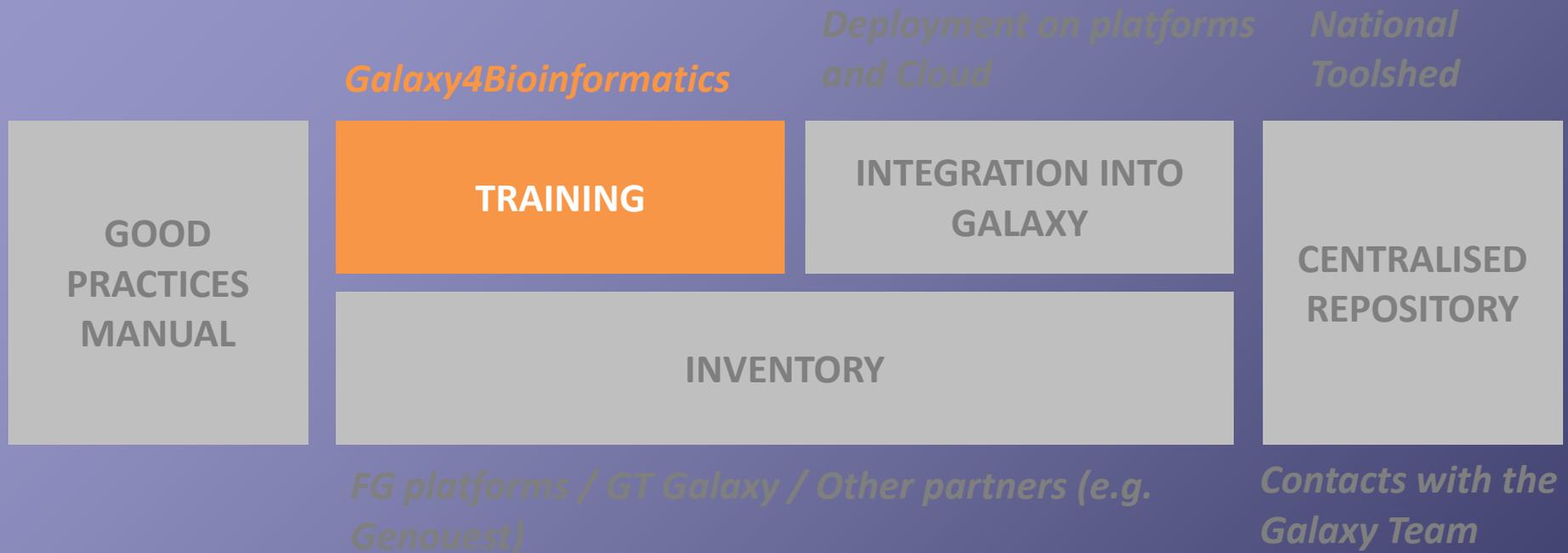
Source : GT Galaxy France

Manuel des bonnes pratiques



- Retour d'expérience des premières écoles bioinformatiques de Roscoff
- Rédaction d'un guide de bonnes pratiques (co-rédaction / wiki)

Animations & Formations



Animations

- Journée(s) d'animation GalaxyDay (s) annuelle(s) 2013- 2014 – 2015
→ Retour d'expériences sur Galaxy (info et bio)

- Animations internationales : Galaxy Community Conference
 - Oslo 2013 (poster, présentations)
 - Baltimore 2014 (présentations)
 - Chicago 2012 - Nebula: A web server dedicated to ChIP-seq data analyses

- Les activités des PFs alimentent le GT.
Et le GT Galaxy aliment les PFs.

Ecole thématique Galaxy4Bioinformatiques

Ecole Thématique CNRS :

- Première école en 2014
- 3 autres écoles en 2015 (Nantes, France Génomique, Sud)



Objectifs :

- Former à l'intégration d'outils + Développements internes
- Renforcer la structuration de la communauté émergente
- 12 Intervenants issus des plateformes de l'IFB (GT Galaxy) -> 20 formés.

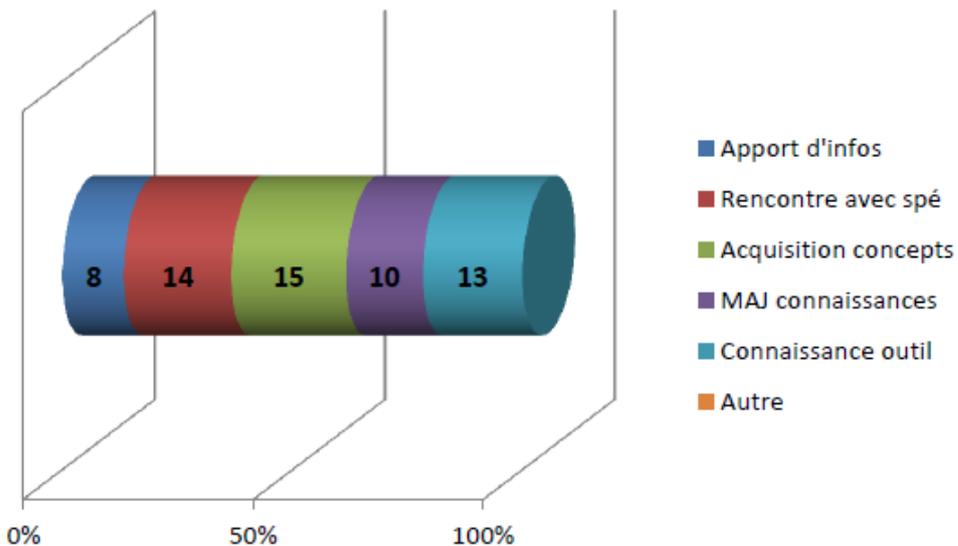
Programme : architecture, ToolShed, API.

Web (supports de cours, programme) : <http://galaxy4bioinformatics.sb-roscoff.fr/>

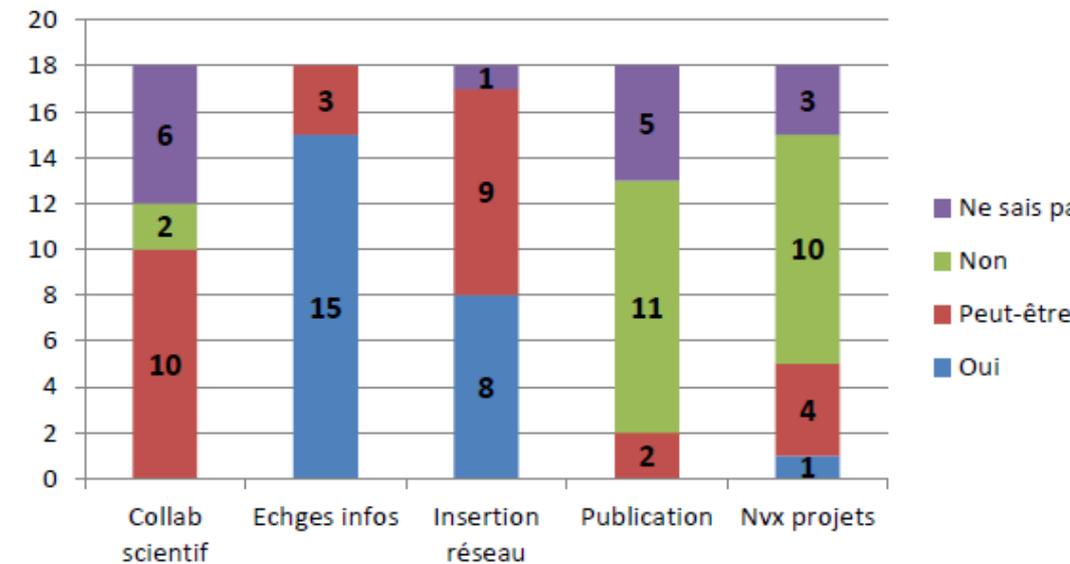
Enquête de satisfaction

Nombre de stagiaires : 20

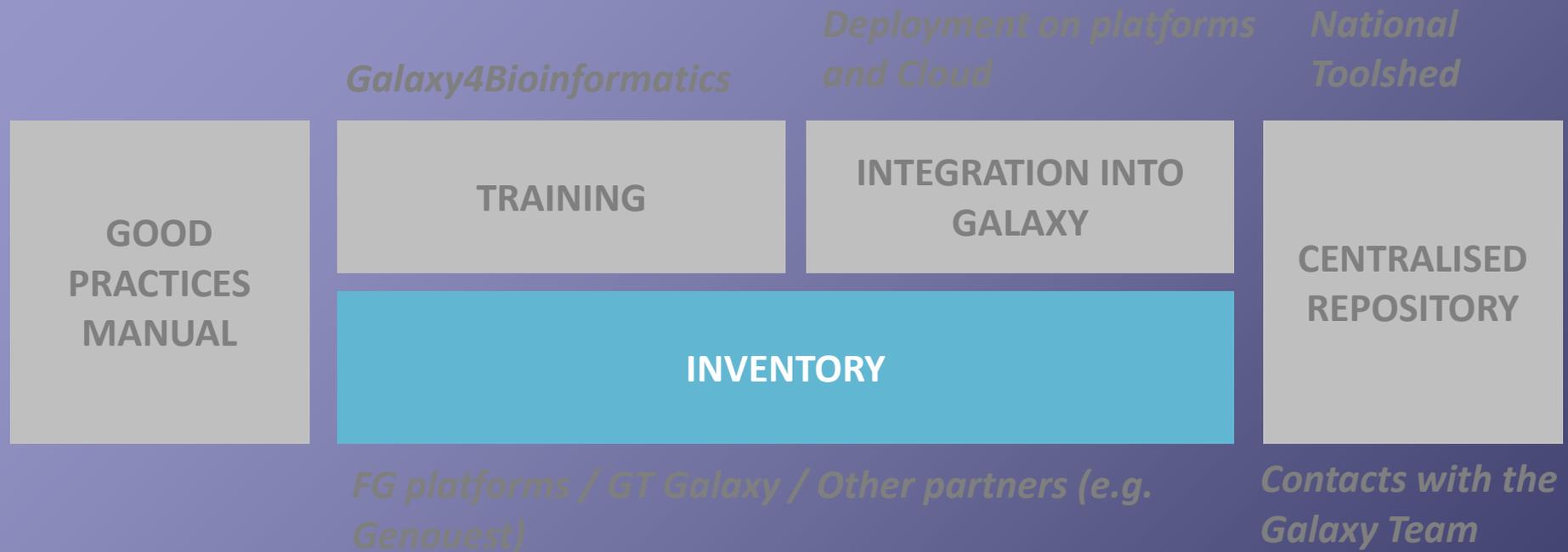
Diriez-vous que cette école a apporté à ses participants :



Pensez-vous que vos échanges conduiront à :

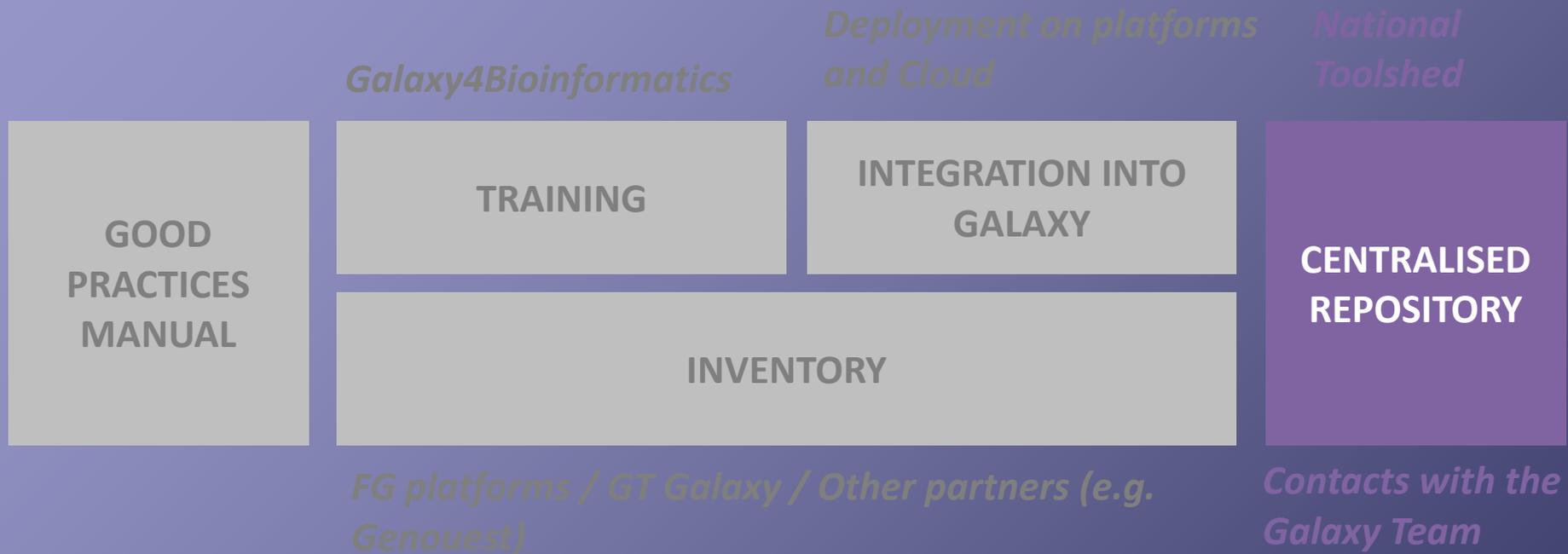


Principales activités du GT Galaxy



- Elaboration d'un inventaire des plateformes Galaxy françaises
- Réponse groupée aux appels offres IFB => 18 mois CDD.

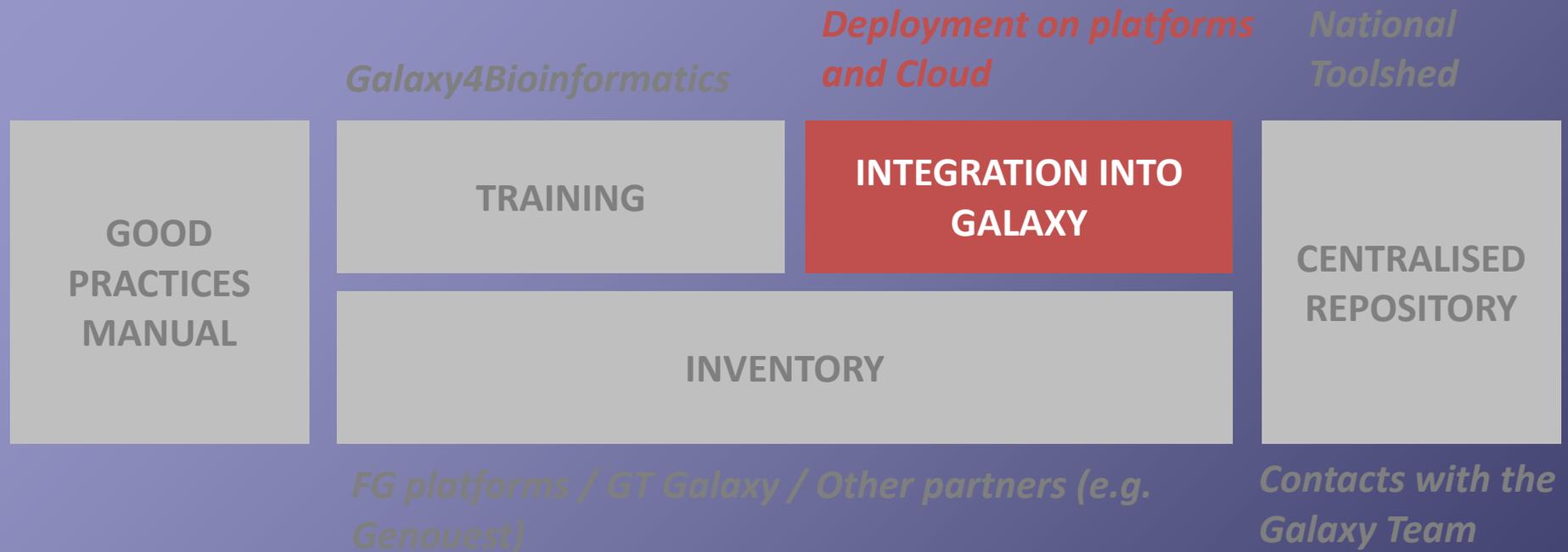
Principales activités du GT Galaxy



ToolShed : International ou French ?

- GT Galaxy échange autour des tests fonctionnels
- GT test le nouveau ToolShed US
→Réunions trimestrielles avec Galaxy project
- Création d'un ToolShed français (en cours) :
 - Visibilité des PF françaises
 - Ajouter des contraintes 'qualité' (bonnes pratiques, tests)

Principales activités du GT Galaxy



- Partenariats pour le développement de wrappers en métabogénomique
- Des idées ? gtgalaxy@groupes.france-bioinformatique.fr

Les évolutions de l'instance Sigenae

Etapes clés de la migration

Quand ?

Octobre 2014

Comment ?

Dans la foulée de la migration de l'infrastructure Genotoul

→ La plus transparente possible pour les utilisateurs

Quoi ?

Récupération de l'ensemble des tools, .loc, paramétrages, banques.

Ajout d'outils et de banques de référence.

SAUF

La base de données...

Amélioration de l'architecture de stockage

Réorganisation de l'infrastructure de stockage et de traitement pour :

- .Protéger au mieux vos données
- .Associer vos données à votre nom
- .Permettre un traitement plus rapide des fichiers par le cluster

Les datasets/logs et runs files étaient archivés dans le `/work/galaxy/`

→ **Dispatching** en fonction du type de fichiers :

- .Fichiers entrants/sortants dans `/galaxydata/`
- .Fichiers tmp, log et scripts de lancement dans des répertoires purgés régulièrement (cron)

Ajout d'un outil de **récupération des datasets dans son `/work/userGalaxy/`**

→ "Libérer" du quota Galaxy

Soumission améliorée

Lancement des jobs au nom de l'utilisateur Galaxy connecté

Pour plus de sécurité:

- .Queue réservée aux traitements Galaxy
- .3 noeuds réservés au jobs Galaxy:
 - .3 noeuds : 032 – 033 – 034
 - .48 slots par noeuds
 - .384 Go de mémoire par noeud

Nouvelle machine:

- .32 Go de mémoire et 8 coeurs (CPUs)
- .Sur la nouvelle infrastructure genovcenter
- .Serveur postgres9 (workstation)
- .Multithreading

Pas de temps de traitement limite de traitement

Possibilité de lancer des jobs plus longs

Paramétrage fin des paramètres du qsub par outil Galaxy

Côté utilisateur Galaxy : Avant - Après la migration

Analyze Data Workflow Shared Data Visualization Admin Help User Welcome smaman, you are working in / Using 13%

Tools Options

Your user name: smaman
Your file path : /work/smaman/

1 - UPLOAD YOUR DATA

Get Data

2 - FILES MANIPULATION

Text Manipulation

Filter and Sort

Join, Subtract and Group

Convert Formats

3 - SEQUENCES MANIPULATION

FASTA manipulation

FASTQ manipulation

SAM/BAM manipulation : Picard (beta)

SAM/BAM manipulation : SAM Tools

4 - MAPPING

BWA - Bowtie

5 - INDEL ET SNP

Indel Analysis

RNA-Seq

GATK Tools (beta)

6 - SRNASEQ

Analyse des miRNA

Annotations

Alignement sur reference





WELCOME ON SIGENAE GALAXY WORKBENCH

Galaxy is a workbench available for biologists from SigenaE Platform. Galaxy objectives are:

- Make bioinfo Linux tools accessible to biologists.
 - Hide the complexity of the infrastructure.
- Allow creation, execution and sharing of workflows.

History Options

TP FastQC 54.0 Mb

8: FastQC data 5.html   

6: GM.fastqsanger   

5: h1.fastqsanger   

4: FastQC data 18.html   

3: FASTQ Summary Statistics on data 18   

2: FASTQ Summary Statistics on data 18   

76 lines, 1 comments
format: tabular, database: ?
Info: 99115 fastq reads were processed.
Based upon quality values and sequence characters, the input data is valid for: sanger
Input ASCII range: '#'(35) - 'C'(67)
Input decimal range: 2 - 34
Epilog : job finished at ven mai 11 10:36:43 CEST 2012

1	2	3	4	5	6
#column	count	min	max	sum	mean
1	99115	2	33	3194703	32.2
2	99115	2	34	3156652	31.8
3	99115	2	34	3145060	31.7
4	99115	2	34	3120431	31.4
5	99115	2	34	3096075	31.2

Côté utilisateur Galaxy : Avant - Après la migration

Sigenae - Welcome smaman1 Workflow Shared Data Visualization Admin Help User Using 354.4 GB

Tools

search tools

YOUR DATA

- [Upload Data](#)
- [Download Data](#)

FILES MANIPULATION

- [Text Manipulation \(e-learning\)](#)
- [Filter and Sort](#)
- [Join, Subtract and Group](#)
- [Convert Formats](#)
- [BED Tools](#)
- [Graph/Display Data](#)

SEQUENCES MANIPULATION

- [FASTA manipulation](#)
- [FASTQ manipulation \(e-learning\)](#)
- [SAM/BAM manipulation : Picard \(beta\)](#)
- [SAM/BAM manipulation: SAMtools \(e-learning\)](#)
- [Fetch Sequences](#)
- [Sequences Queries](#)
- [VCF Tools](#)

SGS MAPPING

- [BWA - Bowtie \(e-learning\)](#)
- [SNP / INDEL](#)
- [NGS: GATK Tools \(beta\)](#)
- [Indel Analysis](#)

WELCOME TO GALAXY WORKBENCH

Galaxy is a workbench available for biologists from Sigenae Platform. Galaxy objectives are:

- Make bioinfo Linux tools accessible to biologists.
- Hide the complexity of the infrastructure.
- Allow creation, execution and sharing of workflows.

NEW BACKUPED DATASET from Sigenae instance of Galaxy before migration

Access to your old version of Galaxy : [Please clic here](#)

Warnings :

- When you access or reload to your Galaxy webpage, please find all your histories saved in the following menu : "User" / "Saved histories".
- Your data are stored in work/ directory. Consequently, BioInfo Genotoul platform reserves the right to purge all files not accessed since 120 days on work/ disk space.

Sigenae support : sigenae-support@listes.inra.fr

If you have some question about Galaxy, please consult your [FAQ](#)

How to cite Galaxy workbench ?

History

Copy of 'star- sampl2' shared by 'nazzouzi@toulouse.inra.fr' (active items only) 187.7 GB

78: report download 7 lines
format: txt, database: ?
cp /galaxydata/database/files/015/dataset_15723.dat //work/smaman/GALAXYtoto55//toto; directory : /work/smaman/GALAXYtoto55/ file(s) saved : tata cp /galaxydata/database/files/016/dataset_16565.dat /work/smaman/GALAXYtoto55/tata ;

```
total 0
drwx--x--x 30 smaman BIOINFO 4.0K Jan 13 1
-rw-r--r-- 1 smaman BIOINFO 338 Jan 13 1
lrwxrwxrwx 1 smaman BIOINFO 48 Jan 13 1
-rw-r--r-- 1 smaman BIOINFO 211 Jan 13 1
drwx--x--x 2 smaman BIOINFO 4.0K Jan 13 1
```

72: report download 5 lines
format: txt, database: ?
cp /galaxydata/database/files/015/dataset_15723.dat //work/smaman/GALAXYtoto3//toto; directory : /work/smaman/GALAXYtoto3/ file(s) saved : titi

Côté utilisateurs

Une nouvelle interface web plus **robuste** :

La charge est équilibrée sur plusieurs threads.

Le code fourni par le « Galaxy Project » est à jour.

Augmentation du nombre d'**utilisateurs** :

2012 : ~ 20

2013 : ~ 40

2014 : ~ 75

40 groupes de travail différents (UMR/Departements)

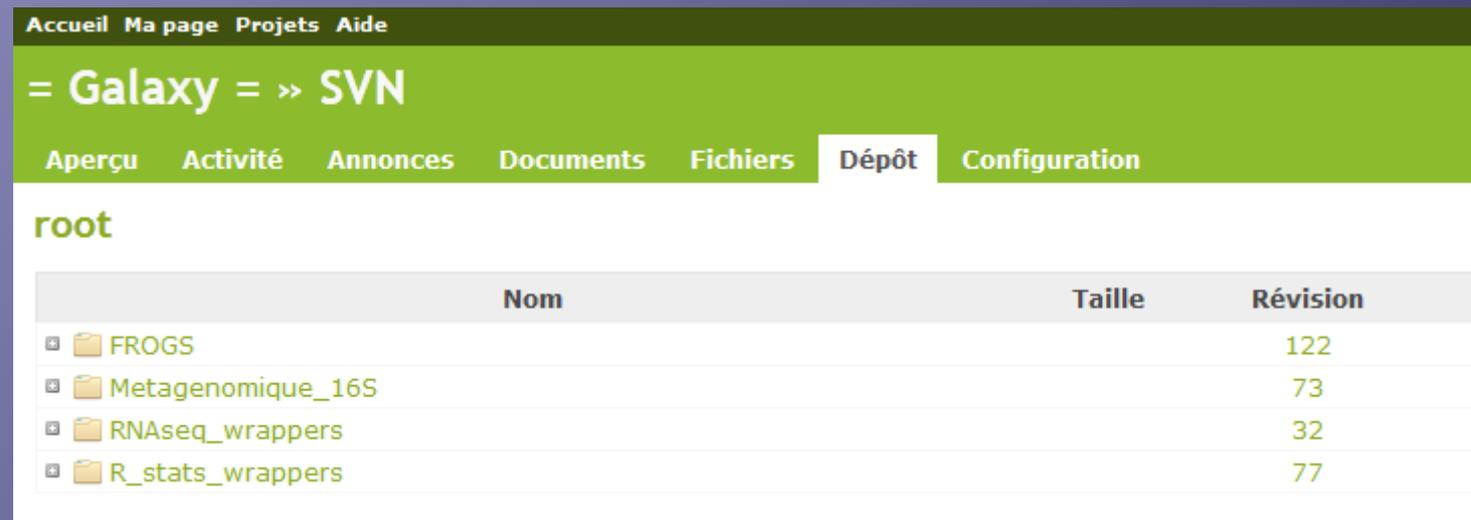
80 nouveaux wrappers

→ **Indépendance** des biologistes sur le traitement de leur données.

<http://galaxy-workbench.toulouse.inra.fr>

Côté groupes de travail

- .Mise à disposition d'une machine de tests
- .Partage de code sous SVN
- .Partage de connaissance (wiki)
- .Action de formations



The screenshot shows a web interface for a Galaxy SVN repository. At the top, there is a navigation bar with links: Accueil, Ma page, Projets, Aide. Below this is a green header with the text "= Galaxy = » SVN". Underneath the header is another navigation bar with links: Aperçu, Activité, Annonces, Documents, Fichiers, Dépôt (highlighted), Configuration. The main content area is titled "root" and contains a table listing the contents of the repository.

	Nom	Taille	Révision
▣	FROGS		122
▣	Metagenomique_16S		73
▣	RNAseq_wrappers		32
▣	R_stats_wrappers		77

Formations sous Galaxy

.Sessions Galaxy sur les PF BioInfo Genotoul PF BioStat

.UPS Master 2 BioInformatique

(Installation, Admin, initiation, traitements bioinformatiques)

.Formation Permanente INRA

(synchronisation avec Jouy-en-Josas)

.Ecole Bioinformatique de Roscoff



The screenshot shows the Aviesan website header with the logo and navigation menu. The main content area features a news article titled "Ecole de bioinformatique : Initiation au traitement des données de génomique obtenues par séquençage à haut débit". The article title is in large, bold, orange text. Below the title, there is a small speaker icon and the word "Ecoute". The breadcrumb navigation shows "Accueil > Toute l'actualité > Ecole de bioinformatique :...".

aviesan
alliance nationale
pour les sciences de la vie et de la santé

Objectifs & Missions

Instituts thématiques multi-organismes

Membres d'Aviesan

Action & Initiatives

Accueil > Toute l'actualité > Ecole de bioinformatique :...

Taille du texte

Ecoute

Ecole de bioinformatique : Initiation au traitement des données de génomique obtenues par séquençage à haut débit

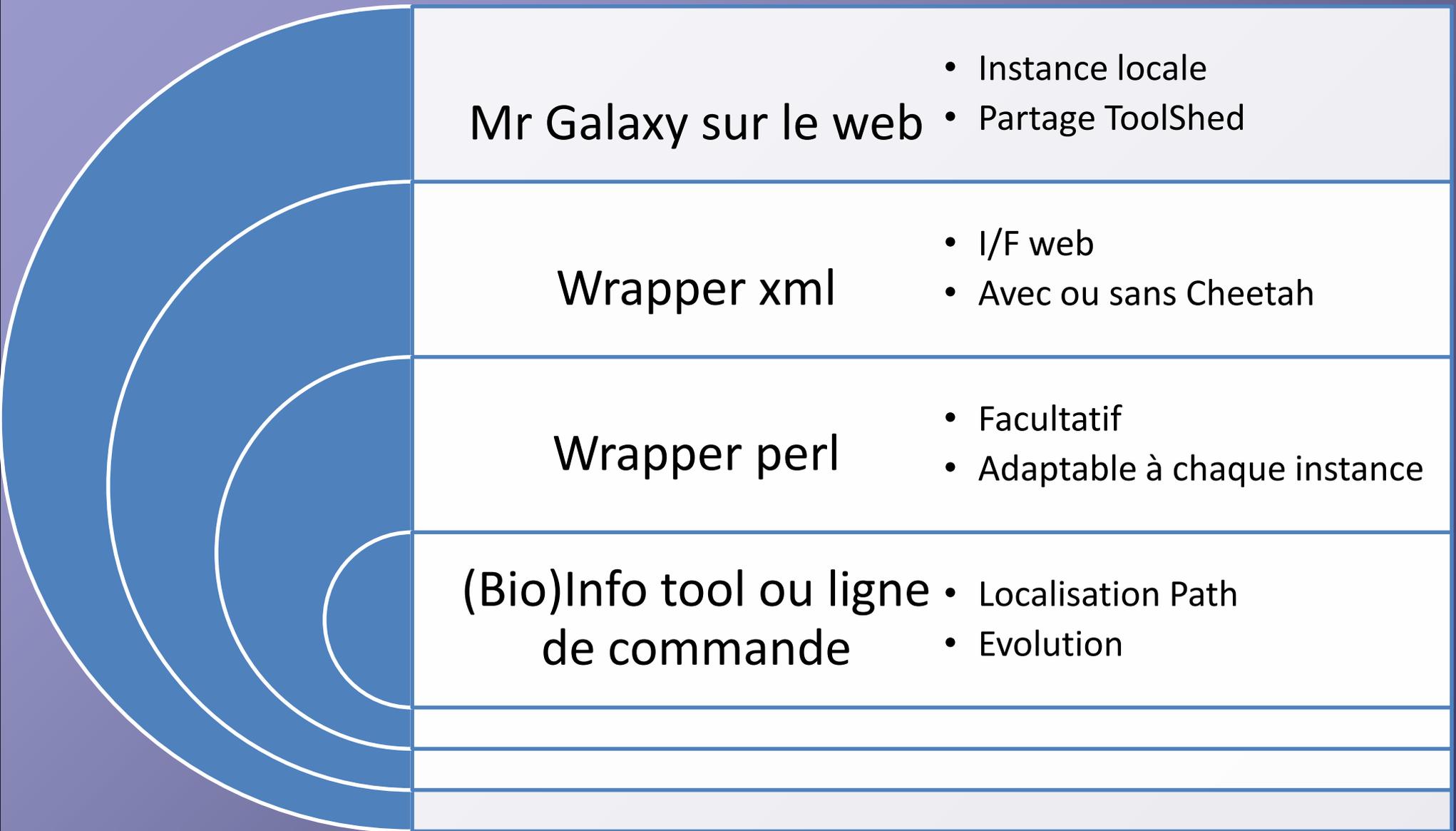
Galaxy training :

- Galaxy : First step
- Galaxy : Reads alignment and SNP calling
- Galaxy : RNAseq alignment and transcripts assemblies
- Galaxy : sRNAseq
- Galaxy : metagenomic : sequence analysis of 16S DNA reads

Actions courantes

- Ajout de génomes de référence (transferts BIOMAJI)
- Ajout de nouveaux wrappers:
 - A coder
 - Issus du ToolShed US
- Paramétrage spécifique du qsub pour chaque tool.
- Assistance / usage de Galaxy mais aussi les outils (choix, paramétrages).
- Gestion des utilisateurs : quotas, jobs management.
- Faire évoluer les wrappers avec les outils bioinfo

Organisation du code en poupées russes



Etapes pour l'ajout d'un nouveau wrapper

1. Vérifier si le wrapper existe déjà ?



Categories of valid repositories

search repository name, description



1. Dans votre instance galaxy ?

2. dans le ToolShed Galaxy Project
(<http://wiki.galaxyproject.org/Tool%20Shed>) ?

3. ou sur le site web de l'outil ?

2. Si le wrapper est dans le ToolShed Galaxy, récupérer le (wget), le paramétrer, puis accéder directement au point 5.

3. Créer le fichier xml dans tools/vosinitiales_myTool/

4. Vérifier les accès aux éventuels outils (bio)informatiques sous jacents.

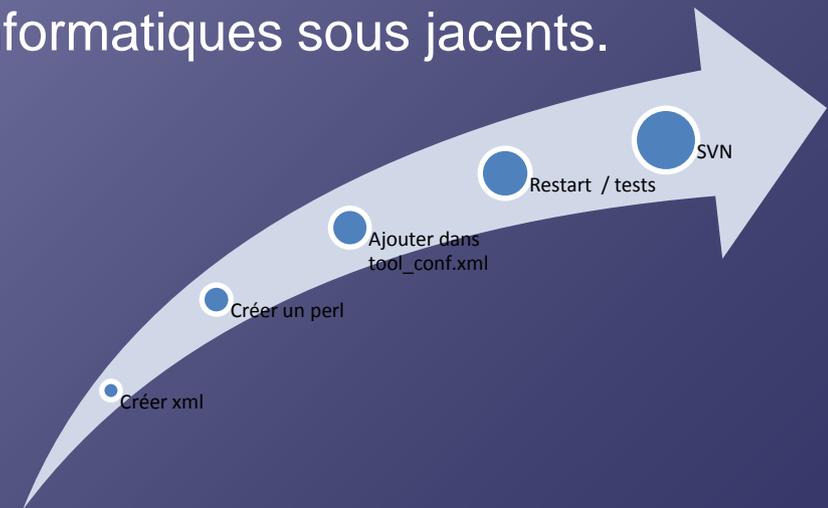
5. Ajouter l'outil dans le tool_conf.xml

6. Configurer en mode exécution

7. Redémarrer la machine pour tester.

8. Tester et corriger -> retour à l'étape 7.

9. Transmettre votre code dans le SVN Galaxy pour un passage en production



Les principales balises du xml

Ce fichier XML est :

- un formulaire de saisie,
- visible depuis l'interface web Galaxy.

```
<tool id = "id_outil" name = "nom_outil " version = "version_outil">  
  <description> description de l'outil </description>  
  <command> perl ou cheeta </command>  
  <inputs>  
    <param ... />  
  </inputs>  
  <outputs>  
    <data ... />  
  </outputs>  
  <help> commentaires, contact, citation </help>  
</tool>
```

```
<tool id="fa_gc_content_1" name="Compute GC content">  
  <description>for each sequence in a file</description>  
  <command interpreter="perl">toolExample.pl $input $output</command>  
  <inputs>  
    <param format="fasta" name="input" type="data" label="Source file"/>  
  </inputs>  
  <outputs>  
    <data format="tabular" name="output" />  
  </outputs>  
  
  <help>  
This tool computes GC content from a FASTA file.  
  </help>  
</tool>
```

Nom interne à Galaxy

Nom du tool affiché sur l'interface web, dans le menu de gauche.

Label affiché sur l'interface web

Help :
reStructuredText
Markup Specification :
<http://docutils.sourceforge.net/docs/ref/rst/restructuredtext.html>

<tool> <description> tags

```
<tool id="sm_mothur_preprocess_V-1-0_date" name="* Nom de l'outil visible dans l'interface galaxy">
  <description>plus de détails</description>
  <command interpreter="perl">sm_wrapper_version_date.pl
    $input
    $param
    "$int "
    $output
  </command>
  <inputs> ... </inputs>
  <outputs> ... </outputs>
  <help> ... </help>
</tool>
```

- L'id du tool doit être unique (integrated_tool_conf.xml)
- Le wrapper est appelé dans la balise <command> du xml.
- Le xml et le wrapper perl sont dans le même répertoire : tools/vosInitiales/fwrapper.pl et file.xml
- Les outils Galaxy sont en anglais.

<command> tag

Exécution directe :

- Ligne de commande : `<command>sed -r '$pattern' $input > $outfile </command>`
- CHEETA (<http://www.cheetahtemplate.org/>):

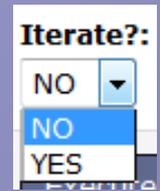
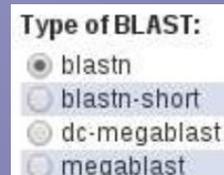
```
<command interpreter="python">gatk_wrapper.py
  --max_jvm_heap_fraction "4"
  --stdout "${output_log}"
  #if str( $ref ) != "None":
    -d "gatk_input"
  #end if <command>
```

Exécution indirecte avec un wrapper :

- un script `ini_tool_wrapper.pl` est appelé dans la balise `<command>`
- Nous appellerons ce script : wrapper.
- Peu importe le langage de ce wrapper. Ils sont tous supportés par Galaxy.
- Il est cependant nécessaire de préciser le langage utilisé dans « interpreter ».

```
<command interpreter="perl">ini_tool_wrapper.pl ... </command>
```

Les principaux composants du tag <input>

	<p>Fichier entrant :</p> <pre><param name = "nom_interne" type = "data" format = "fasta" label = "affichage" /</pre>
	<p>Champs de saisie :</p> <pre><param name = "nom_interne" type = "integer " value = "10.0" label = "affichage"/></pre>
	<p>Liste :</p> <pre><param name = "nom_interne" type = "select " label = « Iterate ?"> <option value = "T">NO</option> <option value = "F">YES</option> </param></pre>
	<p>Boutons radio :</p> <pre><param name = "nom_interne" type = "select " display = "radio" label = "affichage"> <option value = "megablast">megablast</option> <option value = "blastn">blastn</option> <option value = "blastn-short">blastn-short</option> <option value = "dc-megablast">dc-megablast</option> </param></pre>

Les principaux composants du tag <input>

Select a reference genome (

Arabidopsis thaliana ▼

- Arabidopsis thaliana
- Arabidopsis lyrata
- Bos taurus
- Drosophila melanogaster
- Homo sapiens
- Mus musculus
- Rattus norvegicus
- V4_454Scaffolds
- V4_454Scaffolds_filter
- Yeast
- Sus scrofa

1 - XML avec le composant select pour le "Génome de référence "

```
<param name="input_ref_genome" type="select" label="Select a reference genome (if your genome of interest is not listed, please contact Sigenae)">
  <options from_file="mirdeep2_indexes.loc">
    <column name="name" index="0"/>
    <column name="value" index="1"/>
  </options>
</param>
```

2 - mirdeep2_indexes.loc dans tool-data/ :

```
more mirdeep2_indexes.loc
Arabidopsis thaliana /save/galaxy-dev/bank/mirdeep2/Arabidopsis_thaliana
Arabidopsis lyrata /save/galaxy-dev/bank/mirdeep2/ensembl_arabidopsis_lyrata_genome
Bos taurus /save/galaxy-dev/bank/mirdeep2/ensembl_bos_taurus_genome
```

Statistic(s) chosen:

Select All Unselect All

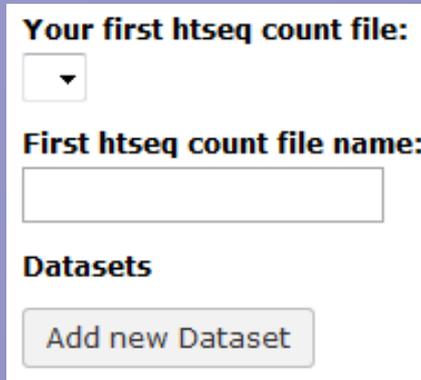
- mean
- sd
- variance
- median
- quartile
- decile

Cases à cocher:

```
<param name="stat_chosen" type="select" display="checkboxes" multiple="True" label="Statistic(s) chosen">
  <option value="mean">mean</option>
  <option value="sd">sd</option>
  <option value="variance">variance</option>
  <option value="median">median</option>
  <option value="quartile">quartile</option>
  <option value="decile">decile</option>
  <validator type="empty_field" message="Please choose a statistic representation" />
</param>
```

Les principaux composants du tag <input>

1 - Répéter un paramètre dans le tag <inputs> du XML :



The screenshot shows a web form with the following elements:

- A dropdown menu labeled "Your first htseq count file:".
- A text input field labeled "First htseq count file name:".
- A section titled "Datasets" containing a button labeled "Add new Dataset".

```
<param format="tab" name="input_htseqcount" type="data" label="Your first htseq count file"/>
<param name="name1" size="20" type="text" value="" label="First htseq count file name"/>
<repeat name="queries" title="Dataset">
  <param name="inputs_count" type="data" format="tab" label="Other htseq count files" />
  <param name="names" size="20" type="text" value="" label="htseq count file name"/>
</repeat>
```

1 - Répéter un paramètre dans le tag <command> du XML :

```
<command interpreter="perl">sm_htseqcount_merge.pl
... #for $q in $queries
    ${q.inputs_count}
    ${q.names}
  #end for
</command>
```

Plus d'infos : <https://wiki.galaxyproject.org>

→ToolConfigSyntax

Les principaux composants du tag <input>

1 – Une condition dans le tag <inputs> du XML :

```
<conditional name="stat_cond">
  <param name="stat" type="select" help="Possible values" label="Stats T or F">
    <option value="T">T</option>
    <option value="F">F</option>
  </param>
  <when value="F" />
  <when value="T">
    <param name="stat_chosen" type="select" display="checkboxes" multiple="True" label="Statistic(s) chosen">
      <option value="mean">mean</option>
      .....
      <validator type="empty_field" message="Please choose a statistic representation" />
    </param>
  </when>
</conditional>
```

```
<conditional name="ploting_cond">
  <param name="ploting" type="select" help="Ploting" label="Ploting T or F">
    <option value="T">T</option>
    <option value="F">F</option>
  </param>
  <when value="F" />
  <when value="T">
    <param name="plot_chosen" type="select" help="" display="checkboxes" multiple="True" label="Plot(s) chosen">
      <option value="boxplot">boxplot</option>
      <option value="histogram">histogram</option>
      <option value="density">density</option>
      <option value="pairsplot">pairsplot</option>
      <option value="MAplot">MAplot</option>
    </param>
  </when>
</conditional>
```

Stats T or F:

F ▼

Possible values

Ploting T or F:

T ▼

Ploting

Plot(s) chosen:

Select All

Unselect All

boxplot

histogram

density

pairsplot

MAplot

2 – Une condition dans le tag <command> du XML :

```
<command interpreter="perl">Graphics_desc.pl
...
  $stat_cond.stat
  $stat_cond.stat_chosen
  $ploting_cond.ploting
  $ploting_cond.plot_chosen
  ....
</command>
```

Les principaux composants du tag <output>

239: Hierarchical classif report

error
An error occurred with this

242: Hierarchical classif report

358 bytes
format: html, database: ?
Group member file NO hclustfun(
count.file = "/work/galaxy-
dev/database/files
/001/dataset_1961.dat",
group.member.file = NULL,
format.image.out = "jpeg",
transformation.method = "none",
sample.clustering = TRUE, se



HTML file

<outputs>

```
<data format = "pdf" name = " result" label " result of ${tool.name} on ${input.name} />
```

```
<filter>pdf is True</filter>
```

```
<!--Ne pas afficher cet output s'il n'est pas généré - - >
```

```
<data name="blast_out" format="tabular" label="BLAST Report">
```

```
<change_format>
```

```
<- - format variable en fonction des conditions d'exécution - - >
```

```
<when input="view_options" value="0" format="txt"/>
```

```
<when input="view_options" value="7" format="blastxml"/>
```

```
<when input="view_options" value="8" format="tabular"/>
```

```
<when input="html_output" value="T" format="html"/>
```

```
</change_format>
```

```
</data>
```

```
</outputs>
```

Les principaux composants du tag <help>

*** Sigcufflinks (version 1.0.0)**

Your accepted hits bam file:

Your gtf file:

G or g ?:
quantitate against reference transcript annot. ▾

Execute

Cufflinks code has been modified in Sigcufflinks by the t

OPTIONS :

- p/--num-threads : number of threads used during ana
- G ou (exclusif) -g :
- G/--GTF quantitate against reference transcript annot
- g/--GTF-guide use reference transcript annotation to

Cufflinks Overview

Cufflinks assembles transcripts, estimates their abunda
reads and assembles the alignments into a parsimonious
many reads support each one. Please cite: Trapnell C,
assembly and abundance estimation from RNA-Seq reve
doi:10.1038/nbt.1621

Know what you are doing

⚠ There is no such thing (yet) as an automated gears

<help>

****Titre en gras****

Pour plus de détails, cliquer ici_..._ ici:
<http://www.google.fr>

.. class:: warningmark
Warning

How to cite

</help>

Wrapper type

Ajouter la licence

```
Copyright (C) 2009 INRA #  
This program is free software: you can redistribute it and/or modify  
it under the terms of the GNU General Public License as published by  
the Free Software Foundation, either version 3 of the License, or  
(at your option) any later version. #  
This program is distributed in the hope that it will be useful,  
but WITHOUT ANY WARRANTY; without even the implied warranty of  
MERCHANTABILITY or FITNESS FOR A PARTICULAR PURPOSE. See the  
GNU General Public License for more details. #  
You should have received a copy of the GNU General Public License  
along with this program. If not, see <http://www.gnu.org/licenses/>
```

Preciser la version de l'outil informatique ou bioinfo sous-jacent

```
<version_command>tophat -version</version_command>
```

Gérer la version grâce au tag <version> du xml

```

#!/usr/bin/perl -w

# usage : perl script.pl <input 1> <output 1>
# Date - Wrapper dans le cadre du projet .... (Pour 'nom du/des biologistes')

use strict;
use File::Basename;
use Getopt::Long;

my $input1;
my $param;
my $output1;
my $selector;
my $conditionalparam1;
my $conditionalparam2;
my $conditionalparam3; ##conditional.param3
my $conditionalparam4; ##conditional.param4

Getopt::Long::Configure( 'no_ignorecase', 'bundling' );
GetOptions (
    'input1=s' => \$input1,
    'param=s' => \$param,
    'output1=s' => \$output1,
    'selector=i' => \$selector,
    'conditionalparam1=i' => \$conditionalparam1,
    'conditionalparam2=i' => \$conditionalparam2,
    'conditionalparam3=s' => \$conditionalparam3,
    'conditionalparam4=s' => \$conditionalparam4
) or die "Usage: Error in command line arguments\n";

my $BIN = '/path/to/your/script/bin.sh';
my $cmd1 = ''; my $cmd2 = ''; my $cmd3 = '';

#Récupérer le numero (unique) de l'output afin, si besoin, de créer un répertoire de travail unique dans /work/galaxy/workspace
my ($nb) = ($output1 =~ /galaxy_dataset_(\d+)\.\S+$/);

#Répertoire de sortie cree par le script, verif des droits d'écriture sur ce repertoire de sortie
`cd /work/galaxy/workspace/; mkdir $nb/; chmod 770 $nb/; cd $nb/;`;
my $dirresults= "/work/galaxy/workspace/.$nb;

#Si besoin, récupération du fichier de configuration avec modification de l extension
`cp $input1 $dirresults/file.txt`;

if ($selector == "1"){
    ..traitements divers...
    #Info pour les biologistes
    print STDOUT "Step 1) - Commenter les traitements effectués : \n\n $cmd1 \n\n ";
}
else
{
    ..traitements divers...
    #Info pour les biologistes
    print STDOUT "Step 2) - Commenter les traitements effectués : \n\n $cmd2 \n\n ";
}

#Si besoin :
#TEST 1 : command ligne on vm-galaxy
#TEST 2 perl Galaxy file : perl script.pl path/to/tests/files/used/for/galaxy/perl/script out1
$cmd3 = "(cd $dirresults; sh $BIN parametes ) >& ./outill.log 2>&1";
system $cmd3;

#Récuperation des fichiers par Galaxy
if (! -e "$dirresults/file.fasta"){print STDERR "file.fasta not found. \n\n";}else{'cp -a $dirresults/file.fasta $output1`;}
..etc...

```

Path d'accès aux outils (bio)informatiques ou scripts wrappés.

Récupération d'un identifiant unique

Création d'un répertoire de travail pour maîtriser le chemin d'accès aux fichiers (ou unlink/move)

Si besoin, modification de l'extension des inputs (.dat par défaut)

Création d'un fichier de logs supplémentaire

Wrappers types disponibles dans le wiki Galaxy



- simple_template.pl : https://forge-dga.jouy.inra.fr/attachments/download/3191/simple_template.pl
- simple_tool_template.xml: https://forge-dga.jouy.inra.fr/attachments/download/3192/simple_tool_template.xml
- template.pl : <https://forge-dga.jouy.inra.fr/attachments/download/3193/template.pl>
- template_podusage.pl : https://forge-dga.jouy.inra.fr/attachments/download/3194/template_podusage.pl

Groupe de travail sur l'instance Sigene de Galaxy

Bonnes pratiques :

- Dev ou modification de wrappers en local sur vm-galaxy-test ---> par les personnes qui souhaitent développer des wrappers (accès actuels : Ibou, Olivier, Frederic, Maria...)
- Puis mise en production (sur vm-galaxy et vm-galaxy-dev) dans un second temps, une fois que les modifications se sont bien déroulées en local (par les admins : Sarah, Patrice, Marie-Stéphane, Didier).

Comment tester vos wrappers sur vm-galaxy-test ?

SVN Galaxy

Voir aussi la section "wrapper type" de ce wiki.

*[[Slides : comment wrapper ?]]

Configuration du wrapper dans le menu Galaxy

Ce fichier décrit la toolbox : le menu de gauche de l'interface Galxy.

```
<?xml version="1.0"?>
<toolbox>

  <label text="Labell" id="feat" />

  <section name="Get Data" id="get_data">
    <tool file="data_source/upload.xml" />
    <tool file="sm_upload/genotoul_upload.xml" />
    <tool file="data_source/ucsc_tablebrowser.xml" />
    <tool file="data_source/ucsc_tablebrowser_test.xml" />
    <tool file="data_source/ucsc_tablebrowser_archaea.xml" />
    <tool file="data_source/microbial_import.xml" />
    <tool file="data_source/biomart.xml" />
    <!--<tool file="sm_save/save.xml"/-->
  </section>

  <section name="Text Manipulation" id="textutil">
    <tool file="stats/column_maker.xml" />
    <tool file="filters/fixValueColumn.xml" />
    <tool file="filters/catWrapper.xml" />
    <tool file="filters/cutWrapper.xml" />
    <tool file="filters/mergeCols.xml" />
    <tool file="filters/convert_characters.xml" />
    <tool file="filters/CreateInterval.xml" />
    <tool file="filters/cutWrapper.xml" />
    <tool file="filters/changeCase.xml" />
    <tool file="filters/pasteWrapper.xml" />
    <tool file="filters/remove_beginning.xml" />
    <tool file="filters/randomlines.xml" />
    <tool file="filters/headWrapper.xml" />
    <tool file="filters/tailWrapper.xml" />
    <tool file="filters/trimmer.xml" />
    <tool file="filters/wc_gnu.xml" />
    <tool file="filters/secure_hash_message_digest.xml" />
    <tool file="new_operations/tables_arithmetic_operations.xml" />
  </section>

</toolbox>
```

Label descriptif

L'outil est placé dans une section.
Le path est relatif.

Configuration du qsub par tool Galaxy

```
<?xml version="1.0"?>
<job_conf>

  <!-- Plateformes d'execution des programmes -->
  <plugins>
    <plugin id="local" type="runner" load="galaxy.jobs.runners.local:LocalJobRunner" workers="4"/>
    <plugin id="sge" type="runner" load="galaxy.jobs.runners.drmaa:DRMAAJobRunner" workers="4"/>
  </plugins>

  <handlers>
    <handler id="main"/>
  </handlers>

  <!-- Chaque destination doit correspondre a une plateforme d'execution-->
  <destinations default="real_user_cluster">
    <destination id="local" runner="local"/>
    <destination id="real_user_cluster" runner="sge">
      <param id="galaxy_external_runjob_script">scripts/drmaa_external_runner.py</param>
      <param id="galaxy_external_killjob_script">scripts/drmaa_external_killer.py</param>
      <param id="galaxy_external_chown_script">scripts/external_chown_script.py</param>
    </destination>
  </destinations>

  <!-- Forcer des outils s'executer sur une seule destination -->
  <tools>
    <tool id="genotoul_upload" destination="local"/>
  </tools>

</job_conf>
```

Plateformes d'exécution

Configuration des destinations

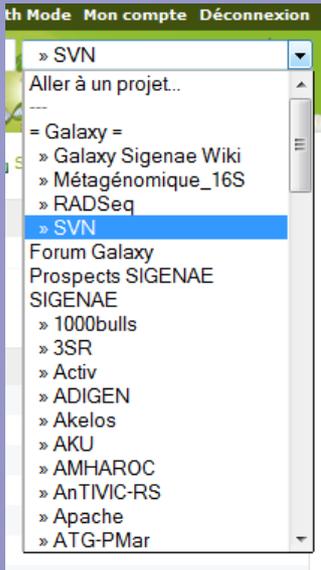
1outil = 1 destination unique

Le fichier job_conf.xml permet de spécifier les paramètres du qsub finement :

```
<destination id="gatk_variant_select_job" runner="drmaa">
  <param id="galaxy_external_runjob_script">scripts/drmaa_external_runner.py</param>
  <param id="galaxy_external_killjob_script">scripts/drmaa_external_killer.py</param>
  <param id="galaxy_external_chown_script">scripts/external_chown_script.py</param>
  <param id="nativeSpecification">-clear -q workq -l mem=10G -l h_vmem=50G -pe parallel_smp 4</param>
</destination>
```

SVN

https://forge-dga.jouy.inra.fr/svn/svn_galaxy_sigenae



SVN Galaxy

Import du svn

```
svn co https://forge-dga.jouy.inra.fr/svn/svn_galaxy_sigenae/
```

Ajout d'un dossier

```
svn import dossier_a_importer https://forge-dga.jouy.inra.fr/svn/svn_galaxy_sigenae/NOUVEAU_DOSSIER_DANS_SVN -m "MESSAGE"
```

Modification des fichiers et propagation au SVN

```
svn add/del fichier  
svn commit
```

Mise à jours

```
svn status  
svn update
```

Interface web d'administration de Galaxy

Administration

Security

- [Manage users](#)
- [Manage groups](#)
- [Manage roles](#)

Data

- [Manage quotas](#)
- [Manage data libraries](#)

Server

- [Tool versions](#)
- [Reload a tool's configuration](#)
- [Profile memory usage](#)
- [Manage jobs](#)

Tool sheds

- [Search and browse tool sheds](#)

Form Definitions

- [Manage form definitions](#)

Sample Tracking

- [Manage sequencers and external services](#)
- [Manage request types](#)
- [Sequencing requests](#)
- [Find samples](#)

Users

Manage users

- Liste des utilisateurs

Manage groups

- Associer un utilisateur à un groupe

Manage roles

- Associer un utilisateur à un rôle

Server

Reload a tool's configuration

- Pour rechargé un outil qui a été commenté dans le tool_conf.xml (préférer le terminal)

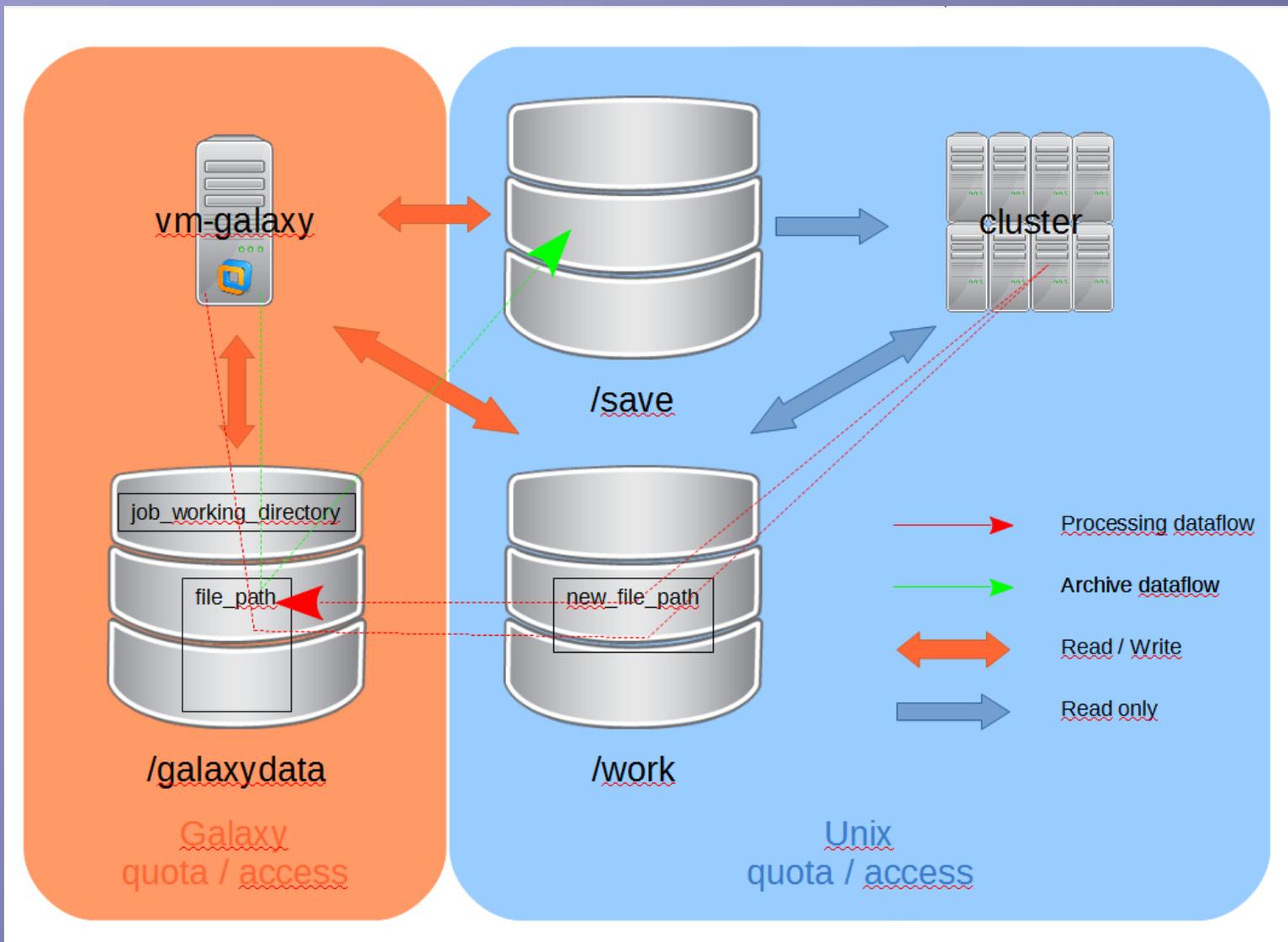
Manage jobs

- Lister les jobs trop longs (en vu d'un kill)

**Merci à toutes celles / tous ceux qui travaillent sur
Galaxy !**

Merci pour votre écoute

sigenae-support@listes.inra.fr



Source : Patrice DEHAIS