



**SIST 2016 : Séries Interopérables et Systèmes de Traitement**  
*29-30 sept. 2016 Montpellier (France)*



# Gestion des données de Cytométrie en Flux

## *Flow Cytometry Data Management*

*Montpellier, le 29 septembre 2016*

Soumaya Lahbib (IE MIO)

Mathilde Dugenne (Doc. MIO), Maurice Libes (OSU), Melilotus Thyssen (Chercheuse MIO),  
Chérif Sammari (Professeur INSTM), Gérald Grégori (Chercheur MIO),  
et Pierre Marrec (Post-Doc MIO)



1. Présentation du projet A\*MIDEX CHROME
2. Introduction au jeu de données de Cytométrie en Flux (CF) automatisée
3. Problématique de la gestion des données de CF
4. Workflow de la gestion des données de CF
5. Consolidation des données
6. Conception, Intégration et restitution
7. Conclusion & perspectives

# 1. Présentation du projet A\*MIDEX CHROME

*Continuous and High Resolution Observation of the Mediterranean Sea*



**CHROME**

Projet CHROME ▾

Equipements ▾

Données et résultats ▾

Contacts



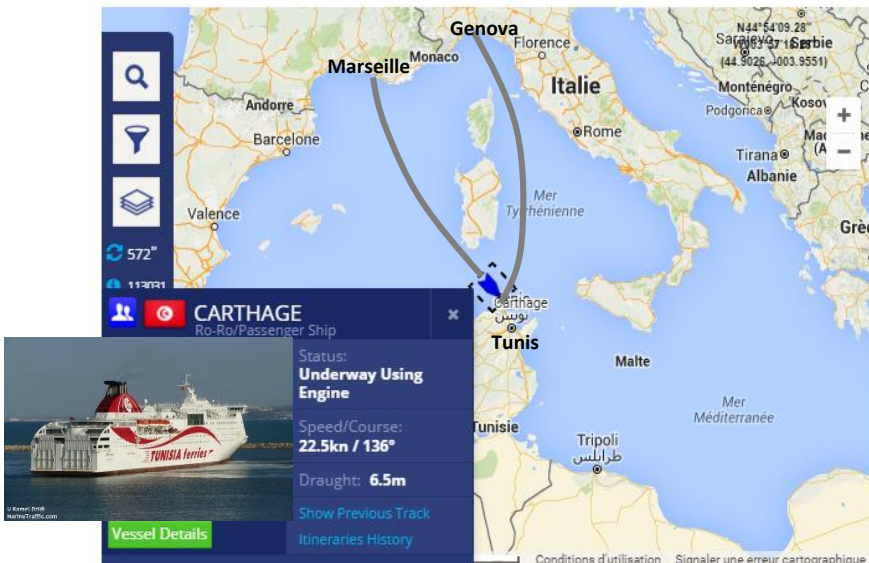
Actualités

Rechercher



# 1. Présentation du projet A\*MIDEX CHROME

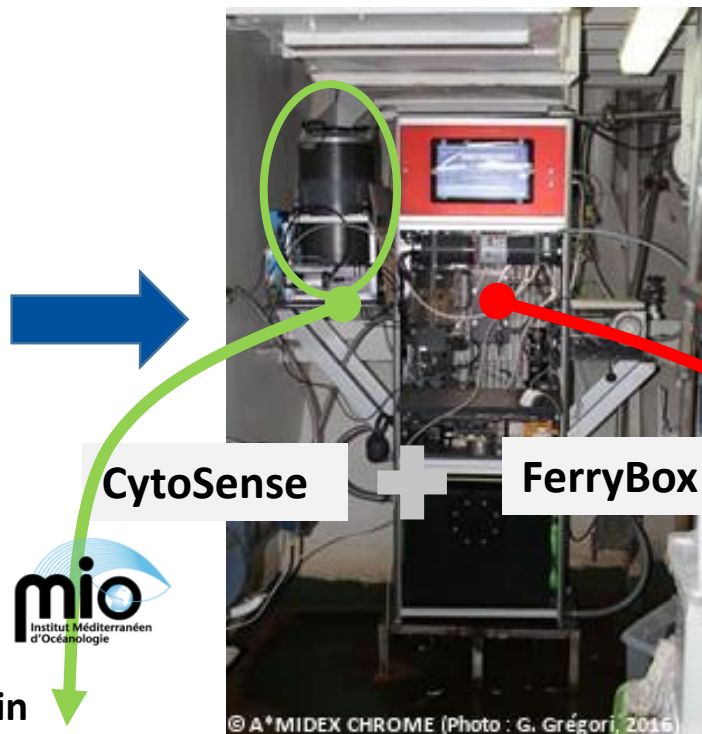
## Pont 1 – C/F CARTHAGE



### C/F CARTHAGE trajectoires

./30min

- Groupes fonctionnels phytoplanctoniques
- Abondance des ces groupes
- Fluorescences/dispersion par cellule
- Estimation de la taille
- Photos des cellules phytoplanctoniques (identification taxonomique > 20 µm)

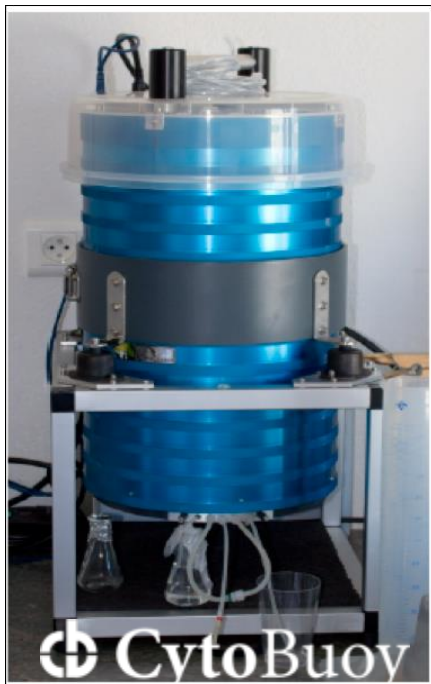


INSTM

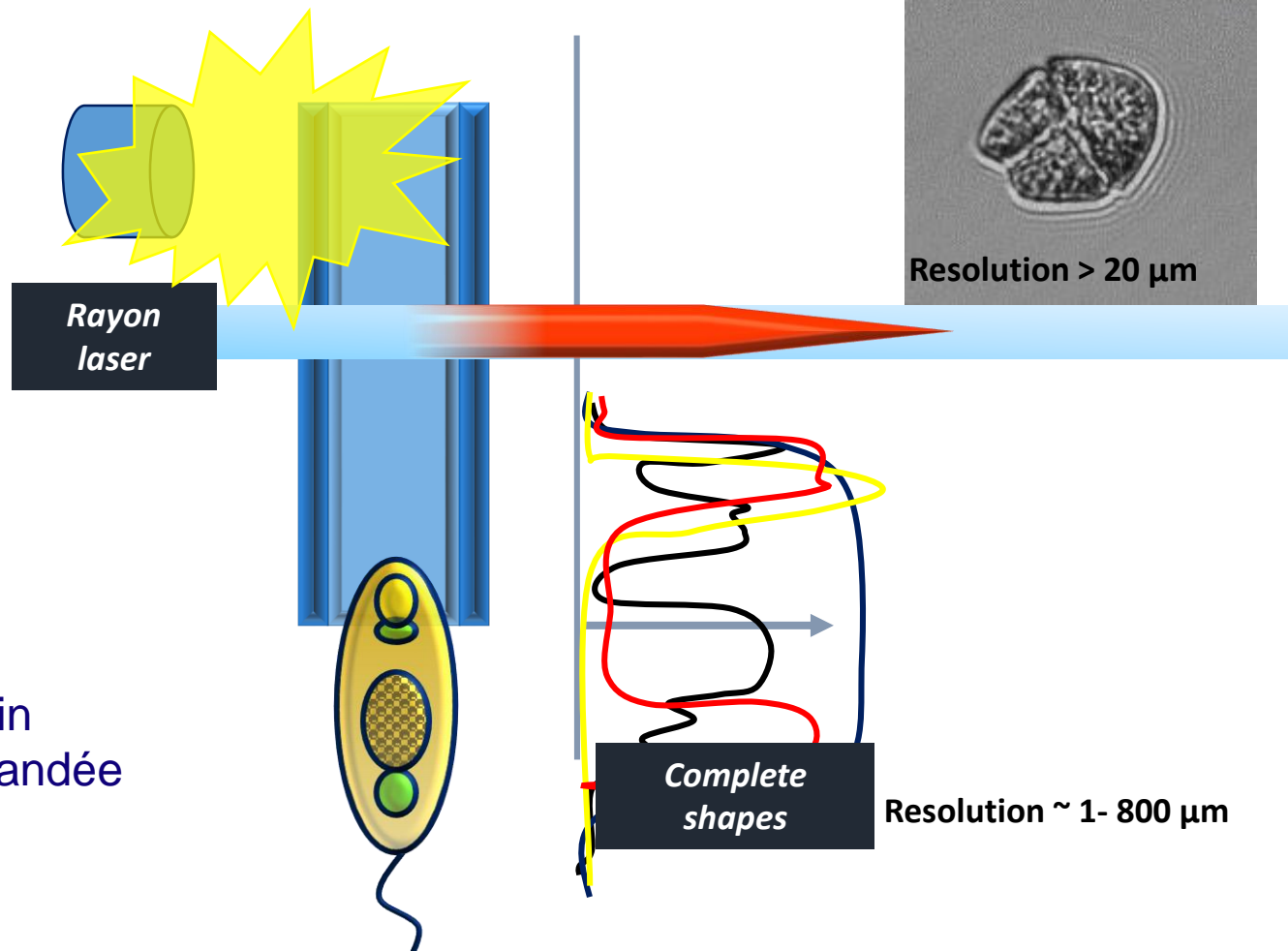
./1min

- Température
- Salinité
- Fluorescence
- Turbidité
- pH
- pCO<sub>2</sub>
- Oxygène

# Nouvelle technologie pour l'étude de la diversité fonctionnelle des groupes phytoplanctoniques à haute résolution spatio-temporelle

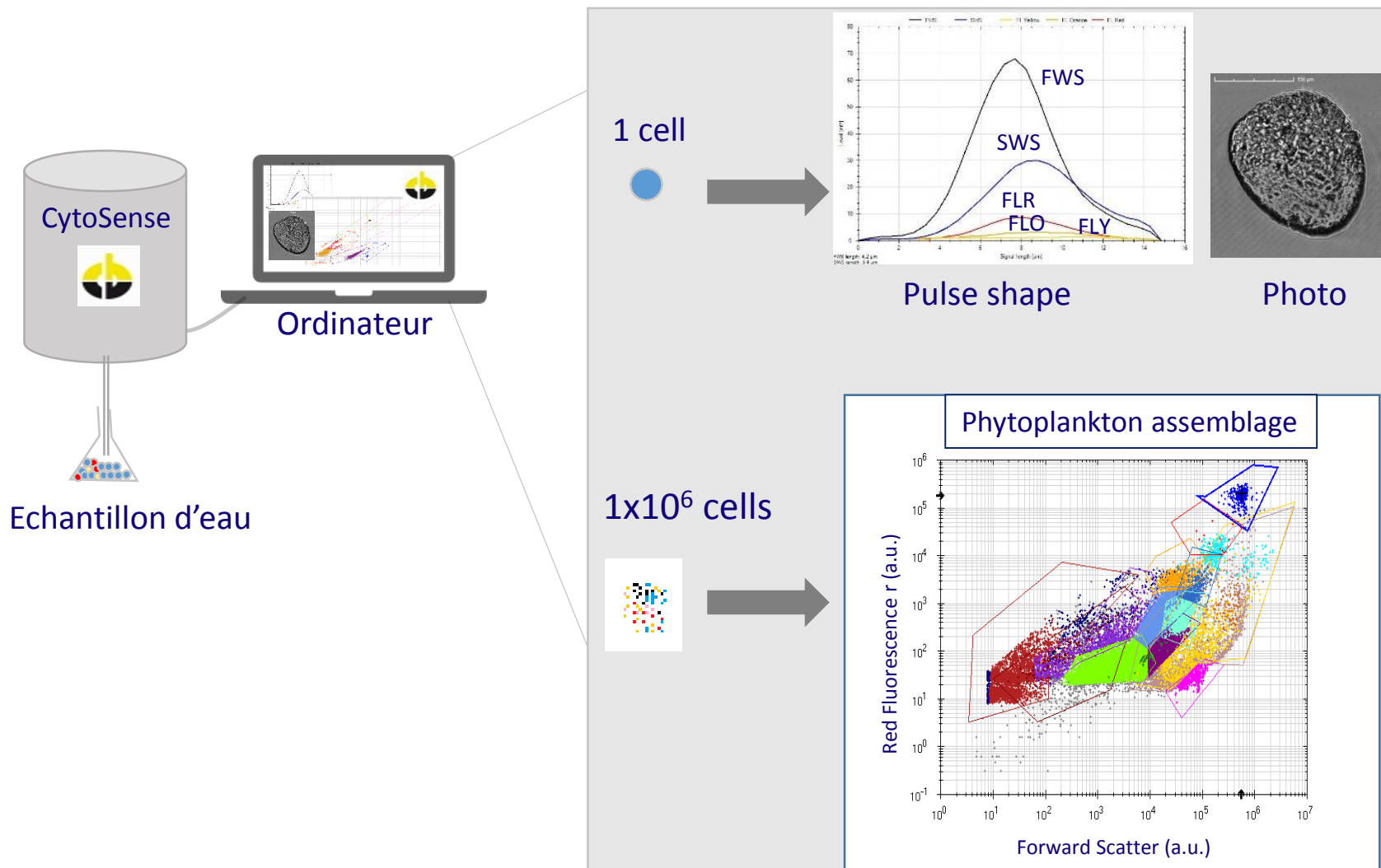


Une analyse (5 cm<sup>3</sup>)/15 min automatisée et télé-commandée





## 2. Introduction au jeu de données de CF automatisée



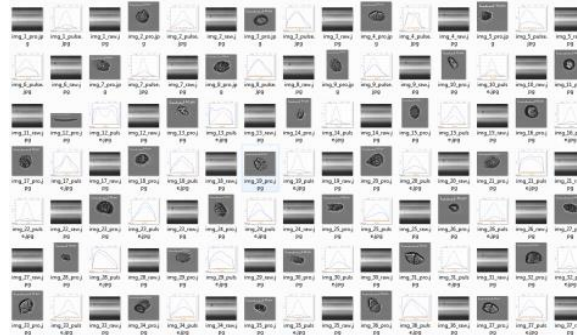
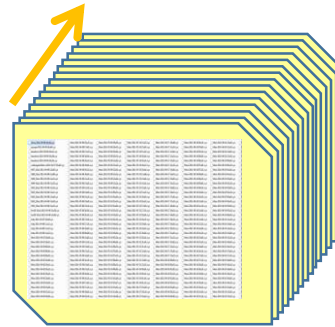
### 3. Problématique de la gestion des données de CF



Mesures par  
le CytoSense



**Classification  
manuelle**



#### Des fichiers statistiques (CSV) :

- Les valeurs Moyenne des propriétés optiques
- Comptage des cellules
- Photos

### 3. Problématique de la gestion des données de CF



1. Grande quantité de données
2. Données mémoirophages
3. Absence de base de données de CFA



La BD CYTOBASE



## 4. Workflow de la gestion des données de CF



Table Données  
Table photo  
Collection  
photos

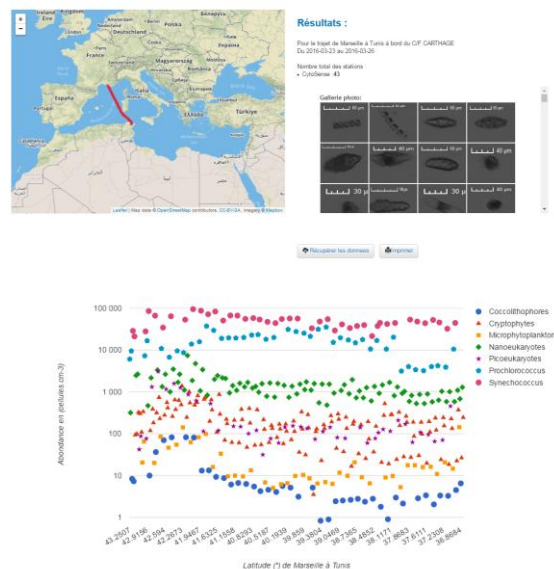


**Acquisition**

**Analyses**

**Consolidation**

**Validation**



**Restitution**

**Intégration**



## 5. Consolidation des données

Fichiers de sorties de Cytoclus	Format	Nombre	Paramètres
<b>Moyenne</b> avgprops_process_xx-xx-xxxx avgprops_process_xx-xx-xxxx avgprops_process_xx-xx-xxxx etc..	csv	1 ou +	Selection set, File name, Date, Time, Total Analyzed Volume [mL], Number of Particles Min. TOF, Max. TOF, Mean TOF, etc...
<b>Comptage</b> counts_process_xx-xx-xxxx counts_process_xx-xx-xxxx counts_process_xx-xx-xxxx counts_process_xx-xx-xxxx etc...	csv	1 ou +	File name, Volume [mL], Trigger chn, Level [mV], Date Particles, Beads 2 mu - Number, Beads 2 mu - pct/Tot Beads 2 mu - N/mL, Microphytoplankton - Number, etc...

# 5. Consolidation : Cytobase Input Processor (M. Dugenne, 2015)

Les fichiers à consolider

Métadonnées

Données

## Cytobase data

Parcourir... 11 fichiers sélectionnés.  
Upload complete

Warning:

Upload successful



Create Inputs folder

Project and samples context   Raw data   Size conversion

<b>Project</b>	<b>Project date</b>	<b>PI</b>	<b>Cytometer ID</b>
<input type="text" value="Enter project name"/>	<input type="text" value="2015-08-18"/>	<input type="text" value="Enter PI name"/>	<input type="text" value="Enter cytometer ID"/>
<b>Station</b>	<b>Depth</b>	<b>Latitude</b>	<b>Longitude</b>
<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>
<b>Filename</b>	<b>Filename model</b>		
<input type="text" value="Filename"/>	<input type="text" value="BERRE_082013_3F_FLR9 2013-12-17 13L"/>		
<b>Samples operator</b>	<b>Standards reference</b>	<b>Clustering method</b>	<b>Observation type</b>
<input type="text" value="Enter name of operator"/>	<input type="text" value="Enter standards beads ref"/>	<input type="text" value="Automated"/>	<input type="text" value="In situ"/>

2013-12-17T14:17:00Z	2013-12-17T14:17:00Z	16,056.60		Synechococcus	BERRE_082013_12S_FLR9 2013-12-17 14u17.cyz	1.99	FL Red	10
2013-12-17T14:17:00Z	2013-12-17T14:17:00Z	16,056.60		Cryptophytes	BERRE_082013_12S_FLR9 2013-12-17 14u17.cyz	1.99	FL Red	10
2013-12-17T14:43:00Z	2013-12-17T14:43:00Z	16,056.61		Beads 2 mu	BERRE_082013_16F_FLR9 2013-12-17 14u43.cyz	2.04	FL Red	10
2013-12-17T14:43:00Z	2013-12-17T14:43:00Z	16,056.61		Microphytoplankton	BERRE_082013_16F_FLR9 2013-12-17 14u43.cyz	2.04	FL Red	10
2013-12-17T14:43:00Z	2013-12-17T14:43:00Z	16,056.61		Picoeukaryotes 2	BERRE_082013_16F_FLR9 2013-12-17 14u43.cyz	2.04	FL Red	10
2013-12-17T14:43:00Z	2013-12-17T14:43:00Z	16,056.61		Picoeukaryotes 1	BERRE_082013_16F_FLR9 2013-12-17 14u43.cyz	2.04	FL Red	10

Please associate each selection set to trigger, PMT's amplification and standardized phytoplankton category

NB: All incompatible entries will be removed

<b>Expert name</b>	<b>Trigger</b>	<b>PMT's amplification</b>	<b>Standardized name</b>
<b>Cluster</b>	<b>Channel/Level</b>	<b>SWS</b>	<b>Cluster</b>
<input type="text" value="Beads 2 mu"/>	<input type="text" value="FL Red 10"/>	<input type="text" value="50"/>	<input type="text" value="Standard beads"/>
		<input type="text" value="50"/>	<input type="text" value="Associate"/>

# 5. Consolidation : Cytobase Input Processor (M. Dugenne, 2015)

Photos

Project and samples context   Raw data   Size conversion   Image-In-Flow pictures   Stations explorer

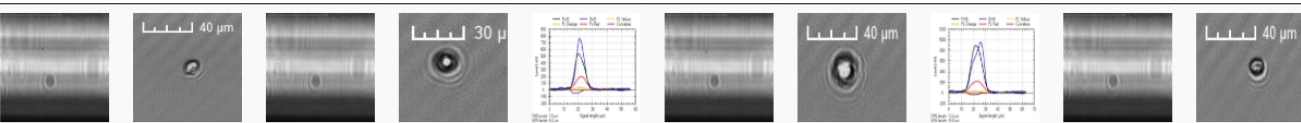
**Add samples pictures**

Select file  
DEWEXL2FLR10 2013-04-05 15u04.cyz

Add pictures  
Choisir les fichiers 15 fichiers   Upload complete

☒ Check all

Show 4 entries   Search:



Project and samples context   Raw data   Size conversion   Image-In-Flow pictures   Stations explorer

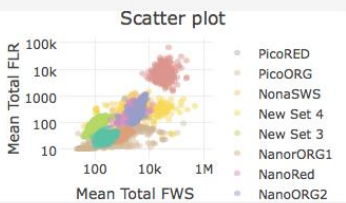
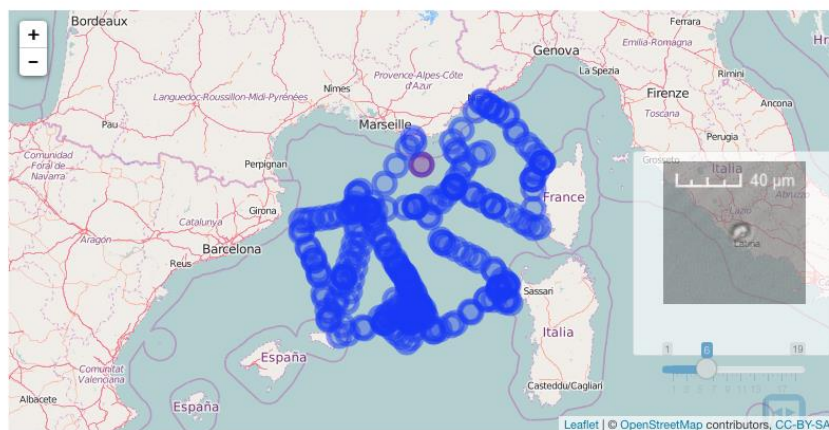
Validation

Select station  
S1

X axis  
Mean Total FWS

Y axis  
Mean Total FLR

Scatter plot

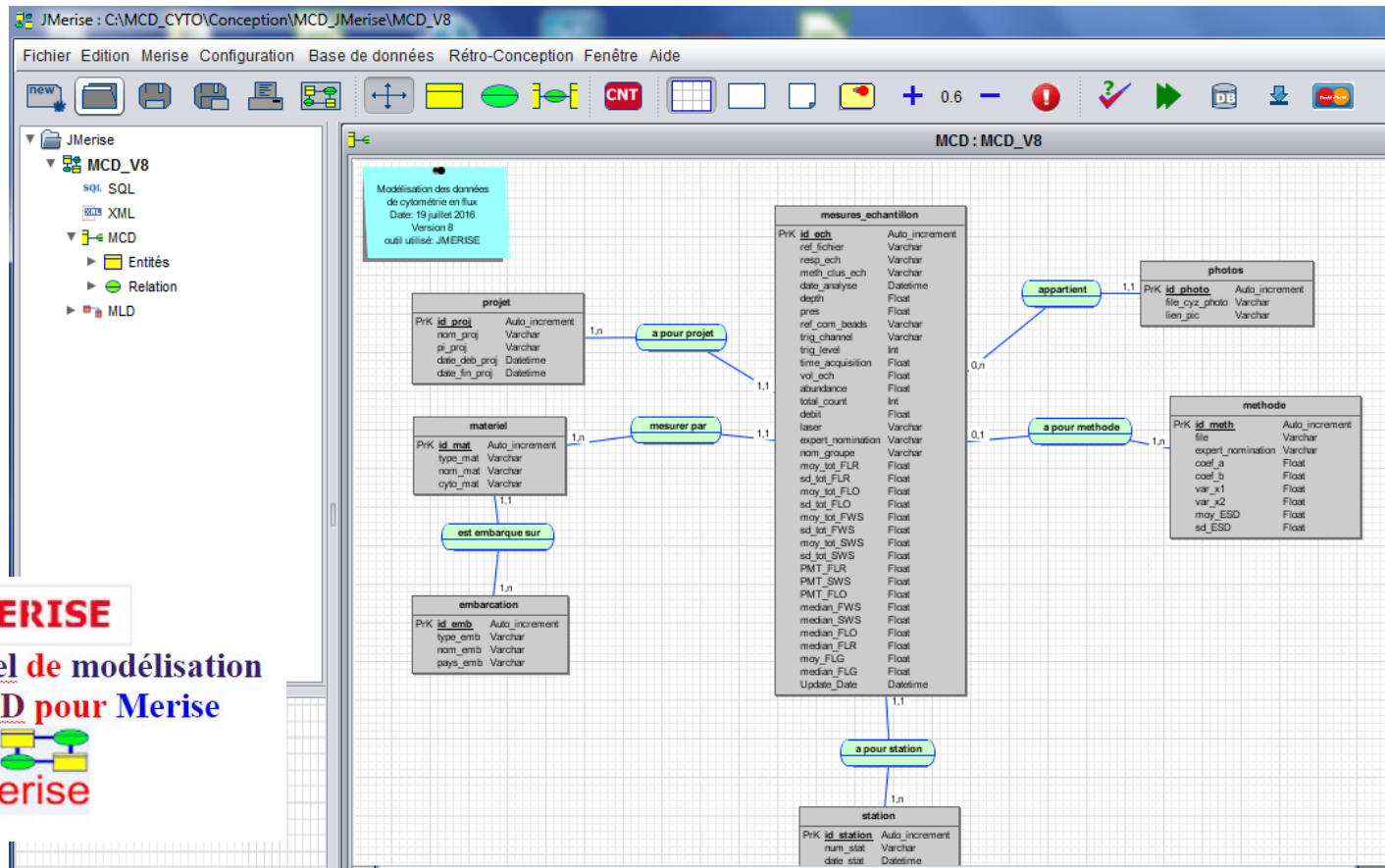



Download Table

Table Données  
Table Photo



## 6. Conception, Intégration et restitution



- Collecte du besoin des utilisateurs
- Fixer les règles de gestion
- Première conception sur papier
- Validation et/ou modification après des réunions

- Choix de Jmerise:
  - Modélisation assez conviviale du MCD :  
Entités et associations
  - Génère le MLD et le script Mysql



## 6. Conception, Intégration et restitution

**phpMyAdmin**

Récentes | Préférences

Nouvelle base de données

chrome\_cytobase

cytobase

Nouvelle table

embarcation

materiel

mesures\_echantillon

methode

photos

projet

station

**MySQL**

Serveur: localhost > Base de données: cytobase

Structure SQL Rechercher Requête Exporter Importer Opérations plus

Table	Action	Lignes	Type	Interclassement	Taille	Perte
embarcation	Afficher Structure Rechercher Insérer Vider Supprimer	8	InnoDB	latin1_swedish_ci	16 Kio	-
materiel	Afficher Structure Rechercher Insérer Vider Supprimer	8	InnoDB	latin1_swedish_ci	16 Kio	-
mesures_echantillon	Afficher Structure Rechercher Insérer Vider Supprimer	17 900	InnoDB	latin1_swedish_ci	4,5 Mio	-
methode	Afficher Structure Rechercher Insérer Vider Supprimer	17 900	InnoDB	latin1_swedish_ci	2,5 Mio	-
photos	Afficher Structure Rechercher Insérer Vider Supprimer	8 801	InnoDB	latin1_swedish_ci	1,6 Mio	-
projet	Afficher Structure Rechercher Insérer Vider Supprimer	8	InnoDB	latin1_swedish_ci	16 Kio	-
station	Afficher Structure Rechercher Insérer Vider Supprimer	1 781	InnoDB	latin1_swedish_ci	240 Kio	-
<b>7 tables</b>	<b>Somme</b>	<b>46 406</b>	<b>InnoDB</b>	<b>latin1_swedish_ci</b>	<b>8,9 Mio</b>	<b>0 0</b>



MySQL Workbench

- Communication avec la base
- Gestion et administration
- Modification et interrogation plus rapide

MySQL Workbench

CytoBase\_Local

File Edit View Database Tools Scripting Help

MySQL Connections

Filter connections

CytoBase\_Local

cytobase-serveur

Shortcuts

MySQL Utilities

Database Migration

MySQL Bug Reporter

Workbench Browser

Planet MySQL

Workbench Forums

Scripting Shell

Models

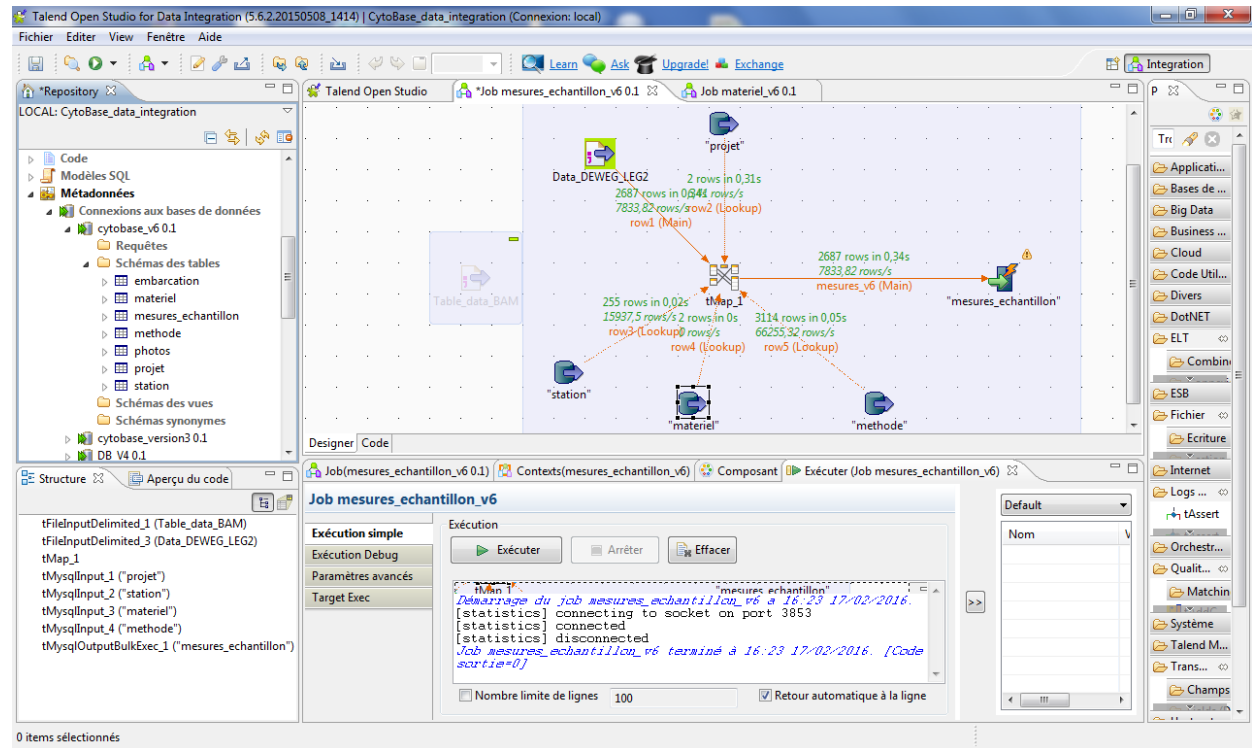
Diagram\_MCD

C:\MCD\_CYTO\Conception

mydb\_cytobase\_local

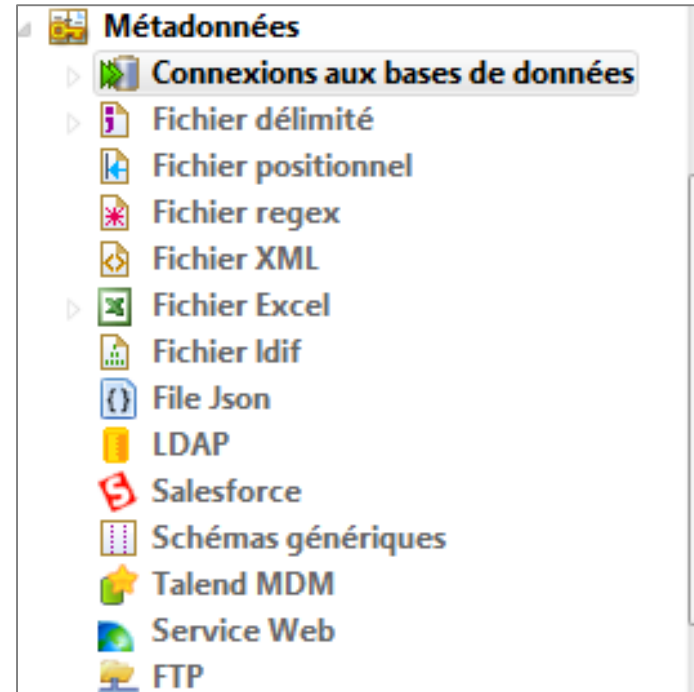
02 Jul 15 11:51

## 6. Conception, Intégration et restitution

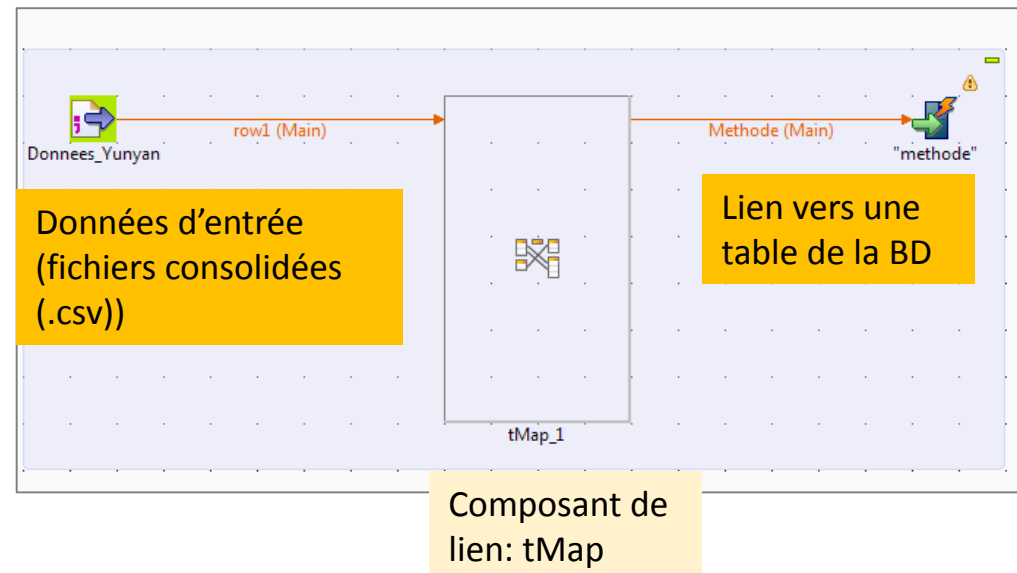


→ Plateforme de développement des traitements ETL (*Extract, Transform, Load*)

→ Connectivité avec les BD, types de fichiers, etc. : BD Mysql



→ Optimisation du temps d'intégration des données consolidées



## 6. Conception, Intégration et restitution

Données d'entrée

Table de la BD

Talend Open Studio for Data Integration - tMap - tMap\_1

Find :

Var

Expression	Type	Variable
TalendDate.parseDate("yyyy-MM-...	Date	var1

Les liens entre les données d'entrées et les champs concernés

Methode

Expression	Column
row1.File	id_meth
row1.Selection_Set	file
row1.Beta_1	expert_nomination
row1.Beta_0	coef_a
row1.Mean_Total_FWS_var1	coef_b
row1.SD_Total_FWS_var2	var_x1
row1.Mean_Length	var_x2
row1.SD_Length	moy_ESD
Var.var1	sd_ESD
	Up_date_Date

Éditeur de Schéma Éditeur d'expression

row1

Colonne	Clé	Type	✓	N..	Modèle date (...)	Length	Precision	Défaut	Comm...
Project	<input type="checkbox"/>	String	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>		14	0		
Project_starting_Date	<input type="checkbox"/>	Date	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	"dd/MM/yyyy"	10	0		
Project_ending_Date	<input type="checkbox"/>	Date	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	"dd/MM/yyyy"	10	0		
PI	<input type="checkbox"/>	String	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>		12	0		
Cytometer ID	<input type="checkbox"/>	String	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>		10	0		

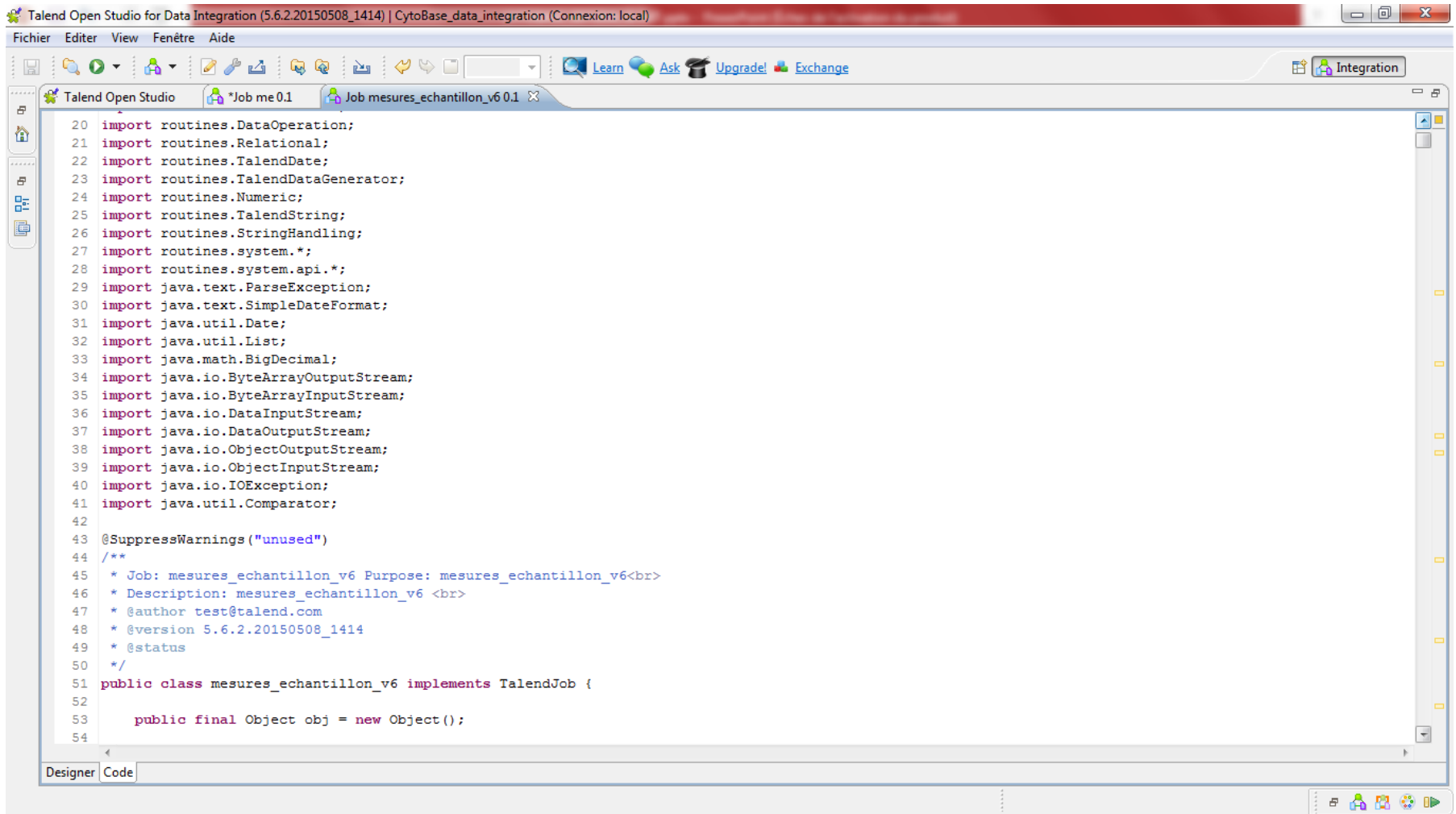
Methode

Colonne	Clé	Type	✓	N..	Modèle date (...)	Length	Precision	Défaut	Comm...
id_meth	<input type="checkbox"/>	int	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>		11	0		
file	<input type="checkbox"/>	String	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>		45	0		
expert_nomination	<input type="checkbox"/>	String	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>		45	0		
coef_a	<input type="checkbox"/>	Float	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>		12	31		
coef_b	<input type="checkbox"/>	Float	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>		12	31		

Appliquer OK Annuler

## 6. Conception, Intégration et restitution

→ Générateur de code en langage Java ou Perl

A screenshot of the Talend Open Studio for Data Integration software. The window title is 'Talend Open Studio for Data Integration (5.6.2.20150508\_1414) | CytoBase\_data\_integration (Connexion: local)'. The menu bar includes 'Fichier', 'Editer', 'View', 'Fenêtre', and 'Aide'. The toolbar contains icons for file operations, development tools, and integration. The main editor shows a Java code file named 'Job mesures\_echantillon\_v6 0.1'. The code includes a series of imports for Talend routines and standard Java classes, followed by a class definition for 'mesures\_echantillon\_v6' that implements 'TalendJob'. The bottom of the window has tabs for 'Designer' and 'Code', with 'Code' currently selected.

```
20 import routines.DataOperation;
21 import routines.Relational;
22 import routines.TalendDate;
23 import routines.TalendDataGenerator;
24 import routines.Numeric;
25 import routines.TalendString;
26 import routines.StringHandling;
27 import routines.system.*;
28 import routines.system.api.*;
29 import java.text.ParseException;
30 import java.text.SimpleDateFormat;
31 import java.util.Date;
32 import java.util.List;
33 import java.math.BigDecimal;
34 import java.io.ByteArrayOutputStream;
35 import java.io.ByteArrayInputStream;
36 import java.io.DataInputStream;
37 import java.io.DataOutputStream;
38 import java.io.ObjectOutputStream;
39 import java.io.ObjectInputStream;
40 import java.io.IOException;
41 import java.util.Comparator;
42
43 @SuppressWarnings("unused")
44 /**
45  * Job: mesures_echantillon_v6 Purpose: mesures_echantillon_v6<br>
46  * Description: mesures_echantillon_v6 <br>
47  * @author test@talend.com
48  * @version 5.6.2.20150508_1414
49  * @status
50  */
51 public class mesures_echantillon_v6 implements TalendJob {
52
53     public final Object obj = new Object();
54 }
```



## 6. Conception, Intégration et développement



*Bibliothèque de fond de carte en javascript*



Google

Developers **dygraphs**

*Génération de graphiques en javascript*

**CHROME** +



<https://chrome.mio.univ-amu.fr/chrome-cytobase/>  
(Accès libre)

<http://www.mio.univ-amu.fr/cytobase/>  
(Accès restreint au MIO)

## 7. Conclusion & perspectives

---

- Un système de mesures automatisées à haute résolution : la FerryBox et le Cytomètre en Flux de type CytoSense.
- Mettre en place une nouvelle méthode de gestion des données de CF (automatisée) et classique.
- Gain de temps dans le traitement et meilleure visualisation et compréhension des données
- Gain considérable d'espace de stockage et de mémoire

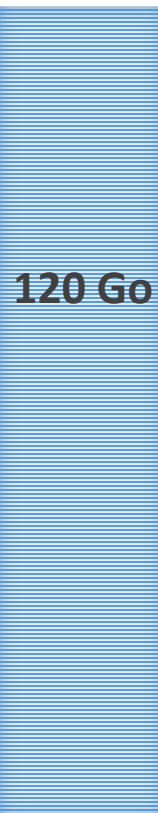
## 7. Conclusion & perspectives

- Cas d'une campagne de 9 jours: 1 échantillonnage/ 20 min

Avant la gestion des données de CF



acquisition

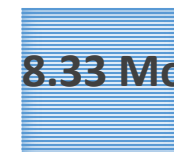


analyses

Après la gestion des données de CF



consolidation



Intégration

## 7. Conclusion & perspectives

---

- La prochaine étape consiste à rendre cette base de données interoperable avec le réseau national (SOMLIT) et international (selon les normes INSPIRE)
- Participation dans le projet EU (SeaDataCloud) pour travailler sur le vocabulaire commun, les standards et les normes à développer pour la CF.

## Remerciements

---

- Le réseau régional des développeurs de Provence ProDev
- Les journées JDEV organisées par le CNRS
- Equipe Osupytheas du M.IO

