



Atelier Technique ODATIS Biologie CES - Cytométrie (en flux)

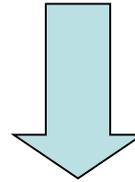
Gérald Grégori (MIO) & Luis Felipe Artigas (LOG)
22-23 Septembre 2020



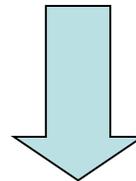
Qu'est-ce que la cytométrie en flux?



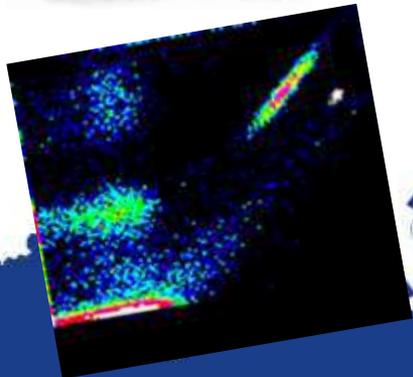
Technique d'origine médicale, destinée au dénombrement et à l'analyse des cellules sanguines



Cellules sanguines = particules unicellulaires en suspension



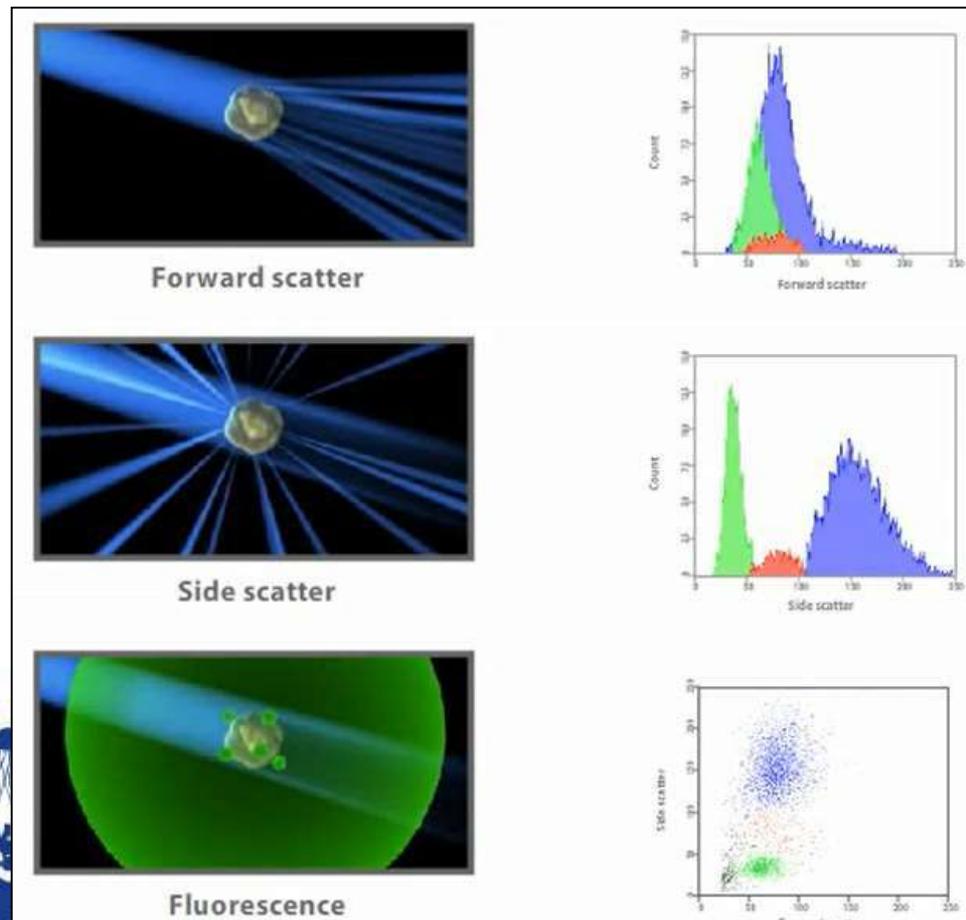
Microorganismes planctoniques



Qu'est-ce que la cytométrie en flux?



Mesure (-métrie) des propriétés optiques de cellules (cyto-) transportées par un liquide vecteur (flux) jusqu'à une source d'excitation lumineuse (le plus souvent un faisceau laser).



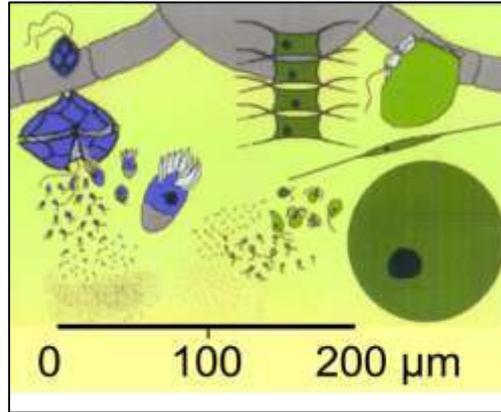
Analyse du phytoplancton



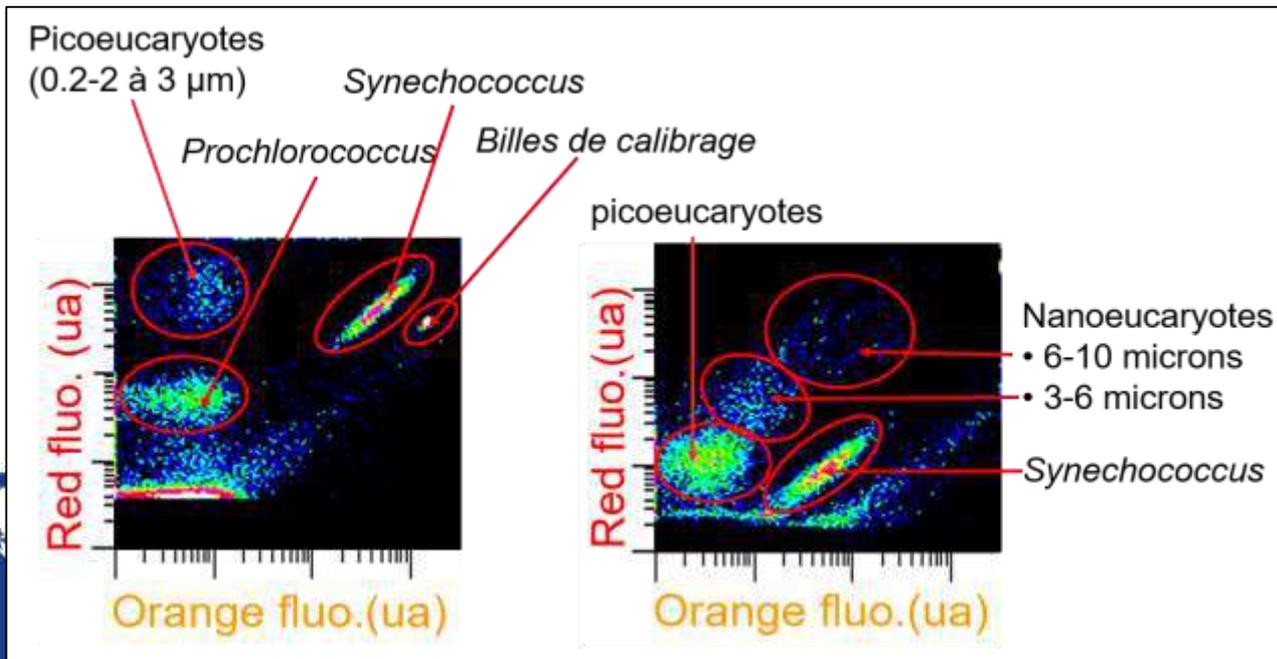
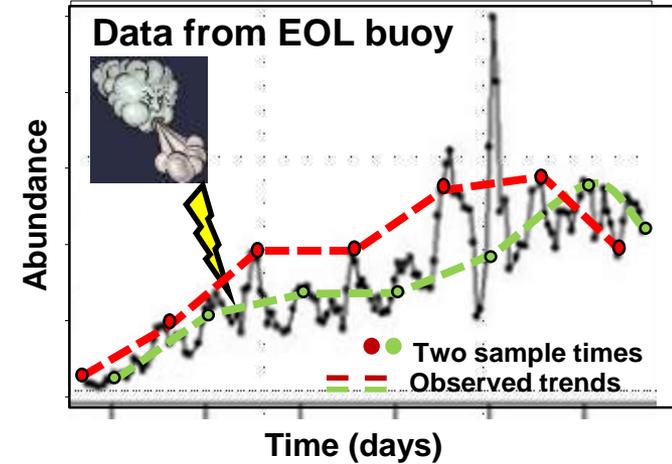
Diversité



Gamme de tailles



Dynamique rapide



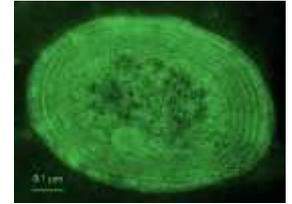
Importantes découvertes réalisées par cytométrie



Prochlorococcus (cyanobactérie)

→ Le plus petit procaryote photoautotrophe et le plus abondant sur la planète

Chisholm, 1988, Nature 334

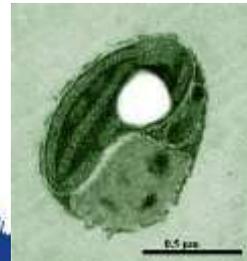
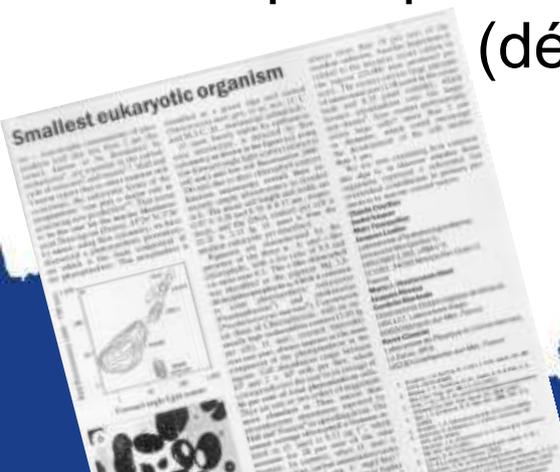


A novel free-living prochlorophyte abundant in the oceanic euphotic zone
Sallie W. Chisholm, Robert J. Olson*, Erik R. Zettler*, Ralf Goerickel†, John B. Waterbury* & Nicholas A. Welschmeyer†
48-425 Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA
* Woods Hole Oceanographic Institution, Woods Hole, Massachusetts 02543, USA
† Harvard University, Cambridge, Massachusetts 02138, USA

Ostreococcus tauri (Chlorophyta, Prasinophyceae)

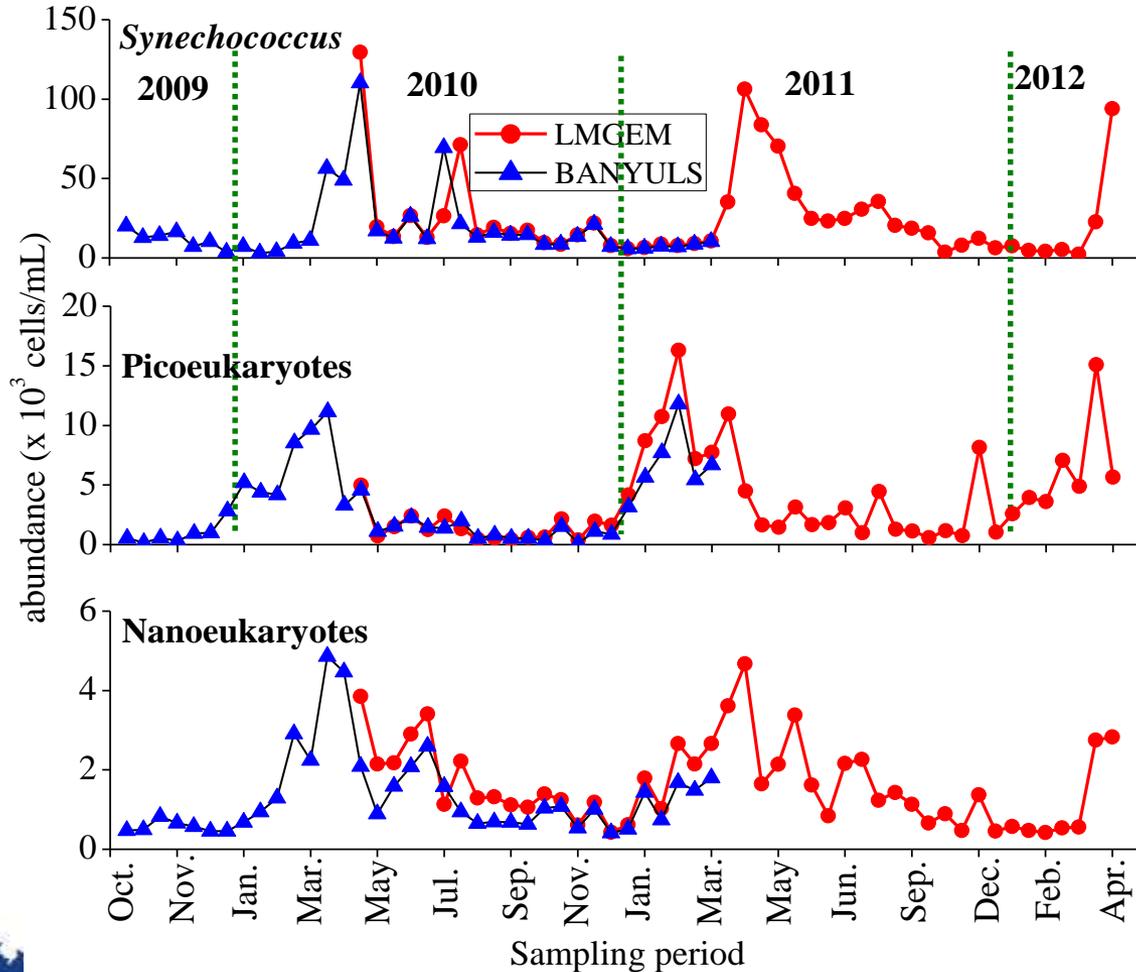
→ Le plus petit eucaryote unicellulaire connu a ce jour (découvert dans l'étang de Thau, France)

Courties et al, 1998, J Phycol 34



Série temporelle du phytoplancton

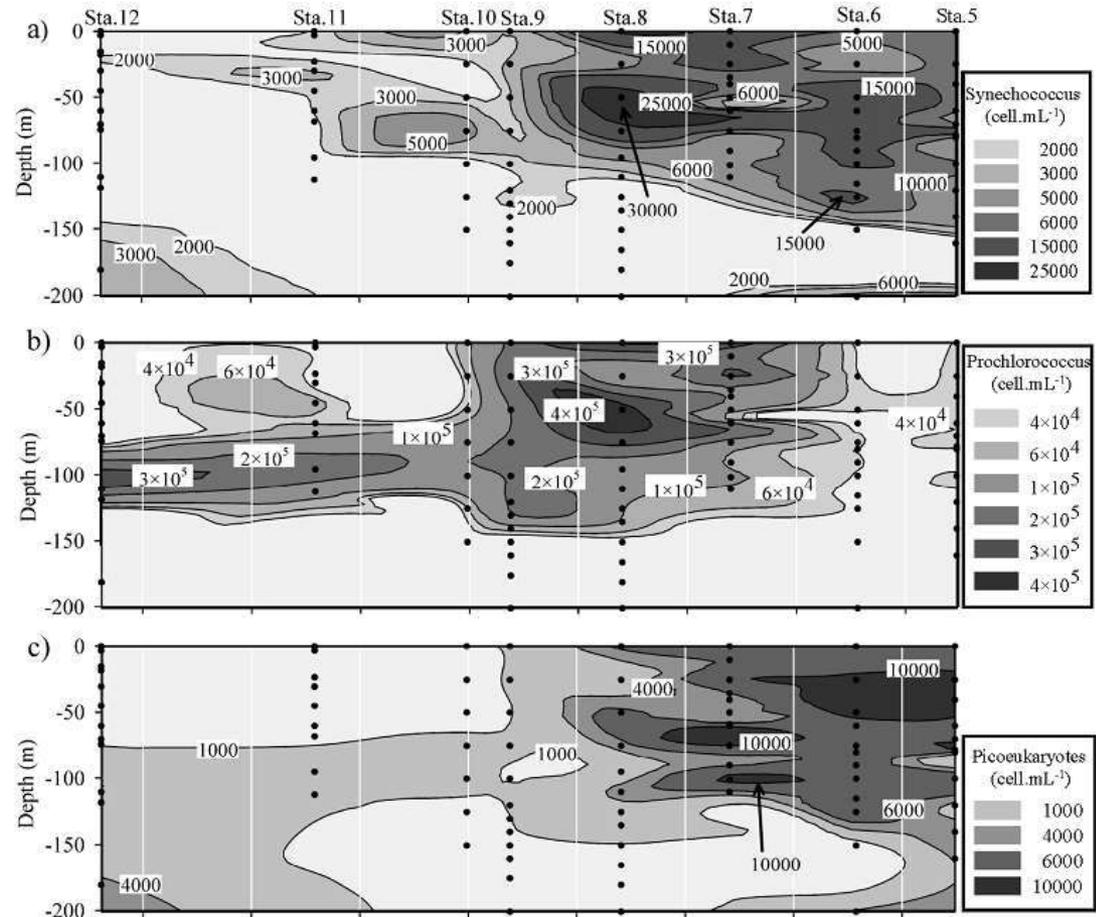
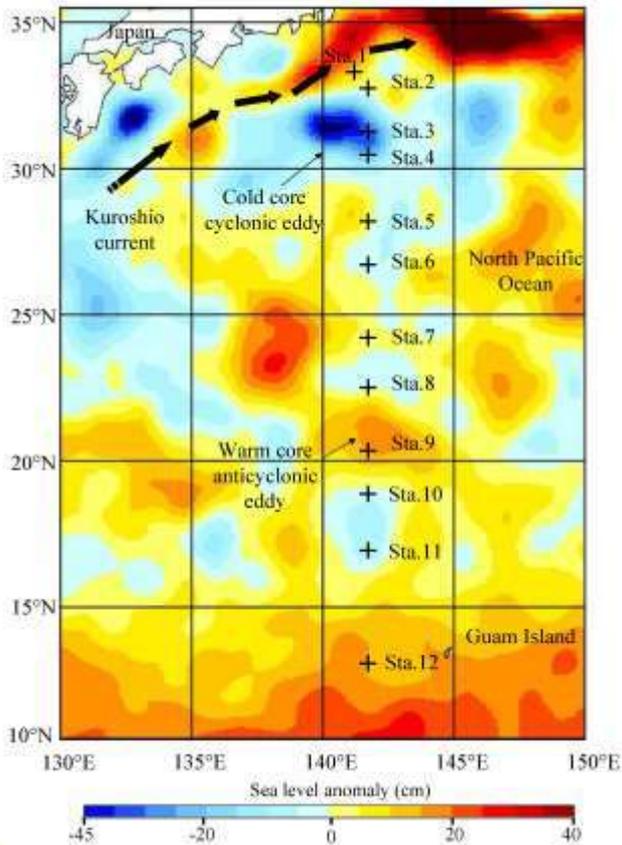
Baie de Marseille – Station SOLEMIO du SNO SOMLIT



Distribution spatiale des abondances du phytoplancton



Tokyo-Palau cruise 2011



Girault et al, 2013. Biogeosciences

Cytométrie automatisée *in situ*



→ Instruments autonomes, déployés *in situ* (navires, bouées, submersibles)

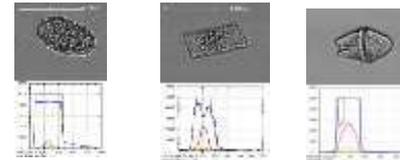
Imaging
FlowCytobot ([IFCB](#))
McLane Research
Laboratories



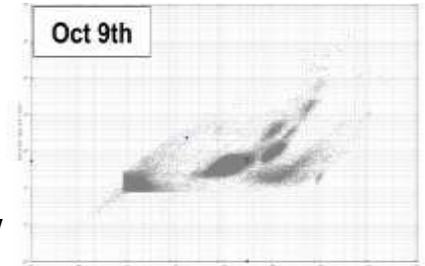
(Sosik et al.)

Images

Cytosense flow
cytometer
Cytobuoy b.v.

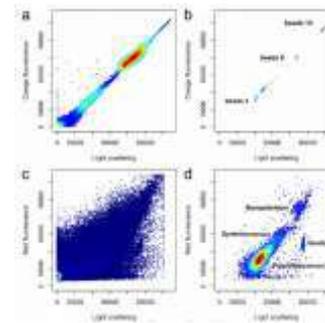


Optical pulses + Image in flow



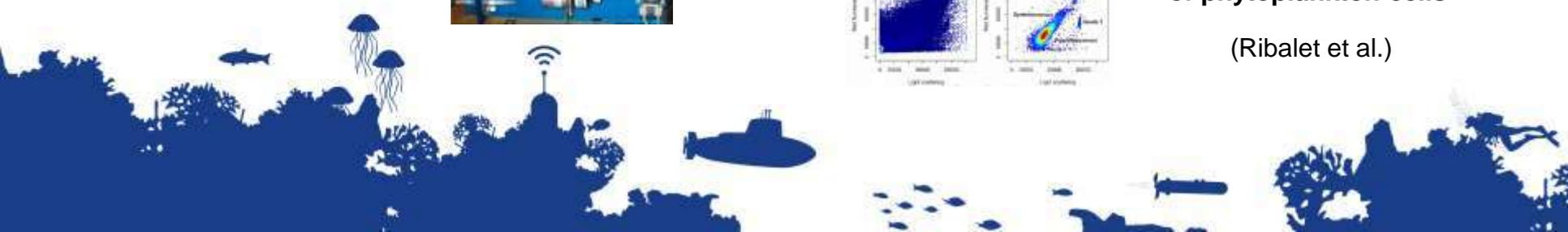
(Thyssen et al.; Dugenne et al.,
Pommati et al.; Artigas et al., etc.)

SeaFlow cytometer
School of Oceanography,
University of Washington



**Continuous, real-time
shipboard observations
of phytoplankton cells**

(Ribalet et al.)



Pourquoi la cytométrie devient populaire en microbiologie marine ?

- **Analyses rapides** (plusieurs dizaines de milliers de cellules analysées par seconde)
 - Permet d'analyser un grand nombre de cellules
 - Résultats statistiques robustes et représentatifs
- **Analyse multiparamétrique à l'échelle individuelle**
- **Données quantitatives** (corrélation avec d'autres données biochimiques)
- **Mesures en temps réel**
- **Classes de taille et abondance cellulaire**
- **Marqueurs uniques:**
 - naturels (chlorophylle, autres pigments autofluorescents)
 - induits (coloration) → fluorochromes (sondes)
- **Tri (post-analyses, cultures)**



Le CES Cytométrie d'ODATIS



- Créé début 2020
- Fédérer les acteurs scientifiques (national, international) autour de la [CMF](#),
- Générer une initiative communautaire pour bancariser, partager et diffuser à l'échelle nationale et internationale les données de cytométrie en flux environnementales (facilement accessibles et réutilisables selon les principes [FAIR](#); qualifiées; répondent aux protocoles d'analyse standardisés reconnus par la communauté).
- Identifier et promouvoir les protocoles (veille technologique) grâce à des échanges d'informations avec la communauté. Identifier les verrous et les outils susceptibles d'améliorer la procédure de génération ou de visualisation de ces données, et leur archivage sur des bases de données interopérables.

- **Coordination et pilotage du [CES](#) Cytométrie en flux** : [Gérald Grégori](#) ([MIO](#), [CNRS](#)),
- **WP Vocabulaire** : Melilotus Thyssen ([MIO](#), [CNRS](#)) et David Pecqueur ([OOB](#), [CNRS](#)),
- **WP Chaîne de traitement des données de cytométrie** : Maurice Libes ([OSU Pytheas](#), [CNRS](#)) et Marc Sourisseau ([IFREMER](#)),
- **WP Bonnes pratiques** : Felipe Artigas ([LOG](#), [ULCO](#)), Christophe Lambert (LEMAR, [CNRS](#))





contact@odatis-ocean.fr
www.odatis-ocean.fr



ODATIS

OCEAN DATA INFORMATION
AND SERVICES

Cytométrie en flux automatisée et accessibilité des données pico-nano- microphytoplancton

CES-ODATIS - SeaDataCloud

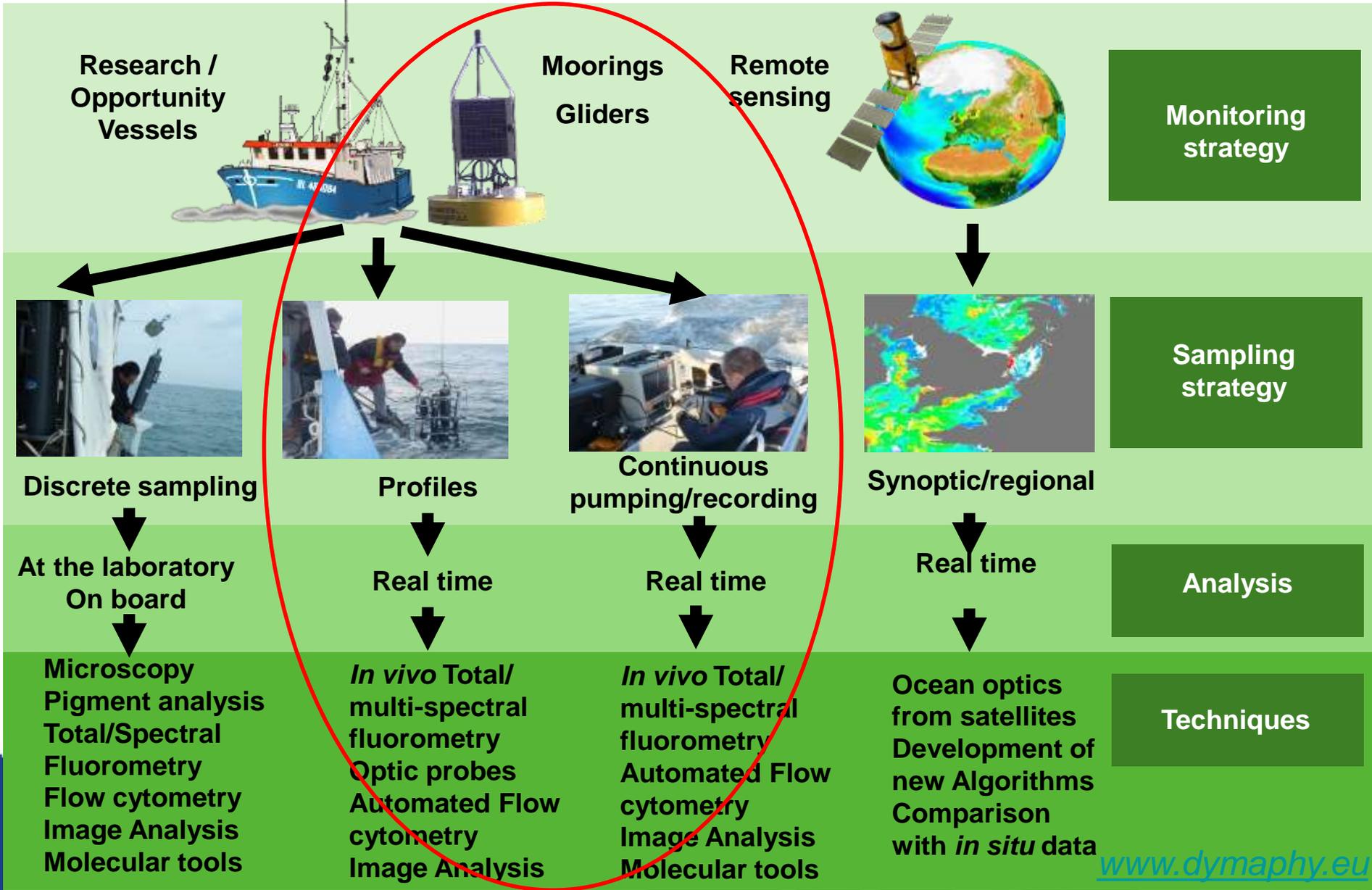
JERICO S3 – JERICO RI

MIO/LOG/IFREMER/VLIZ/RWS/CEFAS

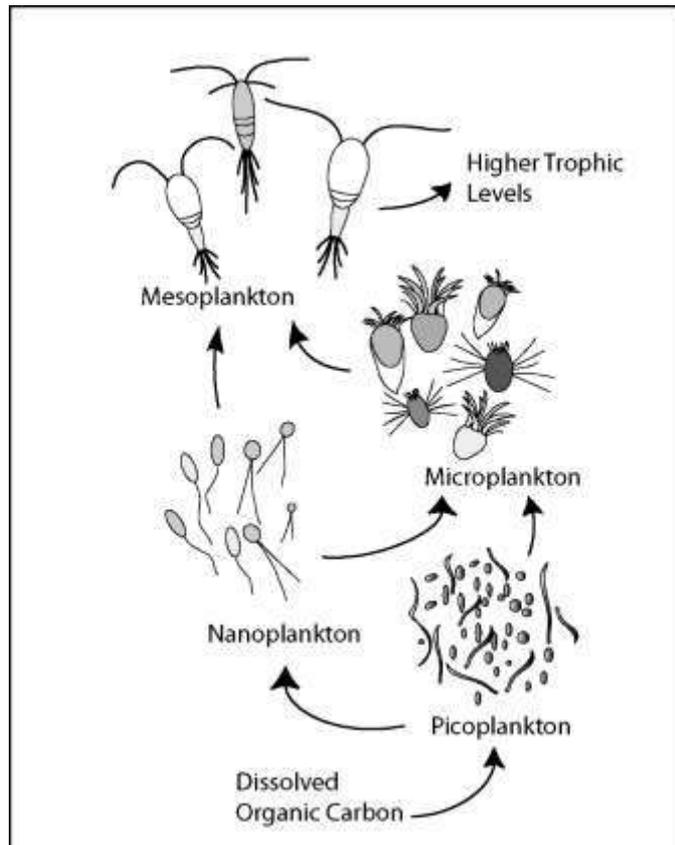
Luis Felipe Artigas (LOG) et Mélilotus Thyssen (MIO)



Méthodes et Approches d'étude du phytoplancton



La diversité fonctionnelle : approfondir la boîte noire



La diversité fonctionnelle (taille, structure, dynamique, abondance, motilité) joue un rôle essentiel dans la productivité et la façon qu'a le carbone organique de circuler dans le réseau trophique et l'exportation via la pompe biologique.

Sa variabilité peut être liée à des changements des conditions environnementales au niveau global, régional ou local.

La cytométrie en flux permet d'accéder à ce niveau de diversité fonctionnelle et son automatiser, permet d'approcher la haute fréquence et la haute résolution spatiale.

Mais le grand nombre de données reste souvent inaccessible à la communauté parce qu'il n'existe pas :

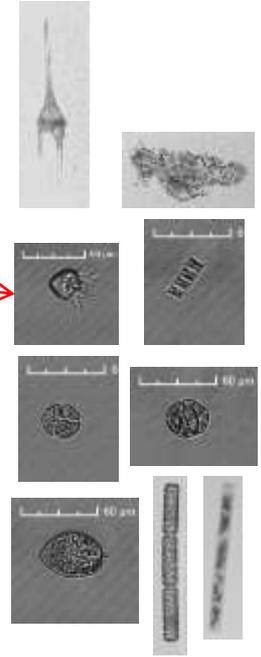
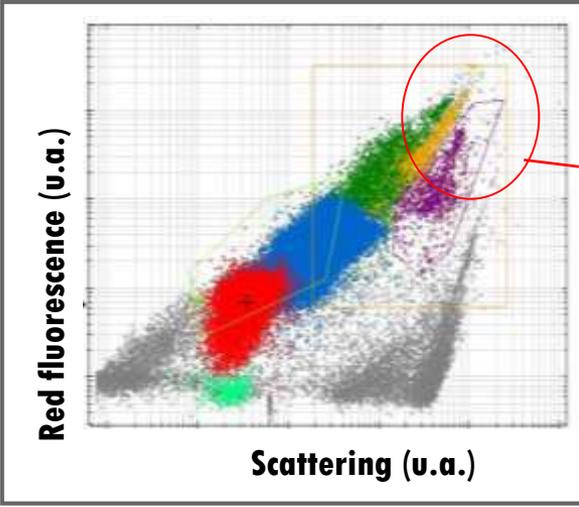
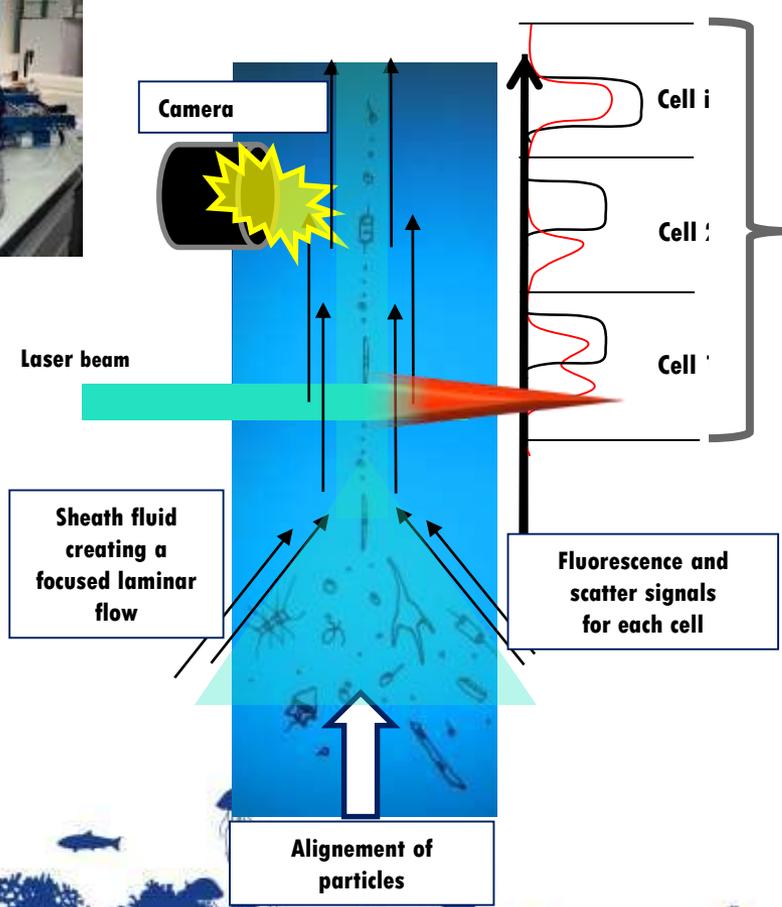
- de vocabulaire commun
- de système interopérable
- de bases de données nationales

/internationales



La cytométrie en flux de type Cytosense permet de détecter et compter, dans un même fichier les plus petites cyanobactéries (*Prochlorococcus*) et les cellules > 200 µm

Analyses automatisées, haute fréquence

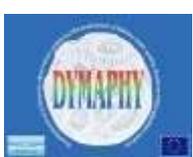


Cytogrammes, profiles and/or pictures describing phytoplankton populations
Up to 7 phytoplankton functional groups resolved :

- Prochlorococcus*
- Synechococcus*
- Picoeukaryotes
- Nanoeukaryotes
- Microphytoplankton
- Cryptophytes like
- Coccolithophoridae like

Vers une interopérabilité des mesures automatisées au niveau des observatoires côtiers à l'international

- Projets recents **d'inter calibration et inter comparaison** des résultats de mesures de différents appareils avec différentes configurations sur des cultures et lors de campagnes en mer internationales (projets INTERREG IV A “2 Mers” DYMAPHY 2010-2014 et H2020 JERICO NEXT 2015-2019)
- Travaux sur un **vocabulaire interopérable** pour permettre aux données issues de la cytométrie en flux d'intégrer **des bases de données marines** (projet SeaData Cloud : accessibilité via les portails de gestion de données marines – 2016-2020)
- Vers une meilleure **interopérabilité** de l'utilisation de la technique de cytométrie en flux automatisée (CES Cytométrie ODATIS, JERICO S3 et autres projets) :
 - définition de **procédures opérationnelles communes** aux utilisateurs de la technique au niveau de la communauté scientifique
 - **automatisation des analyses des signaux** (supervisée, semi-supervisée et non supervisée) en développant outils existants et en explorant d'autres. Proposition **d'accès virtuels** à certains outils
 - avancement dans la **definition des formats et qualification des données** pour la mise en place de flux (**workflows**) alimentant les **bases de données**.



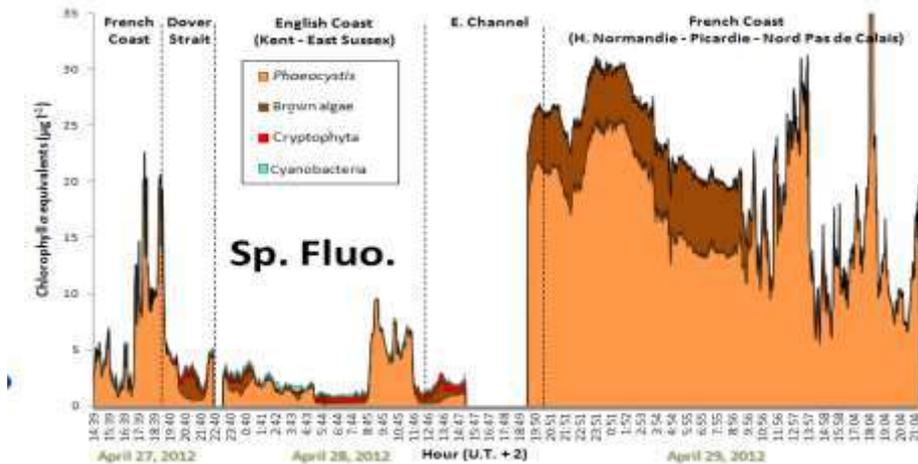
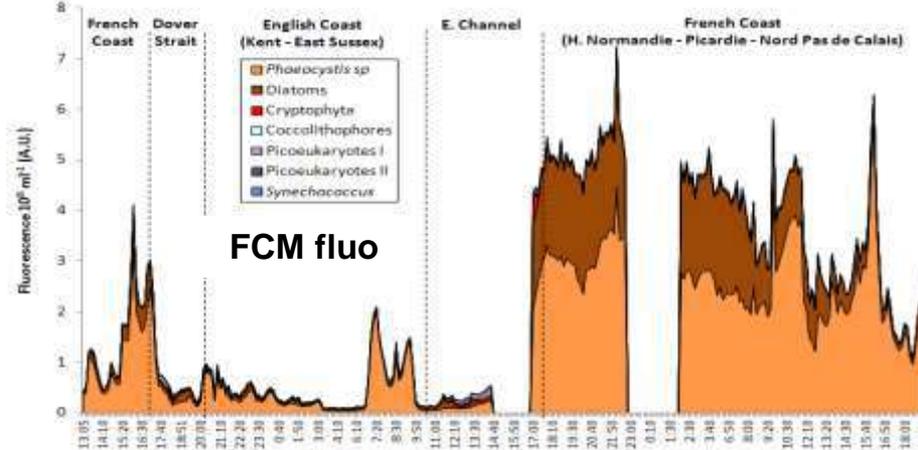
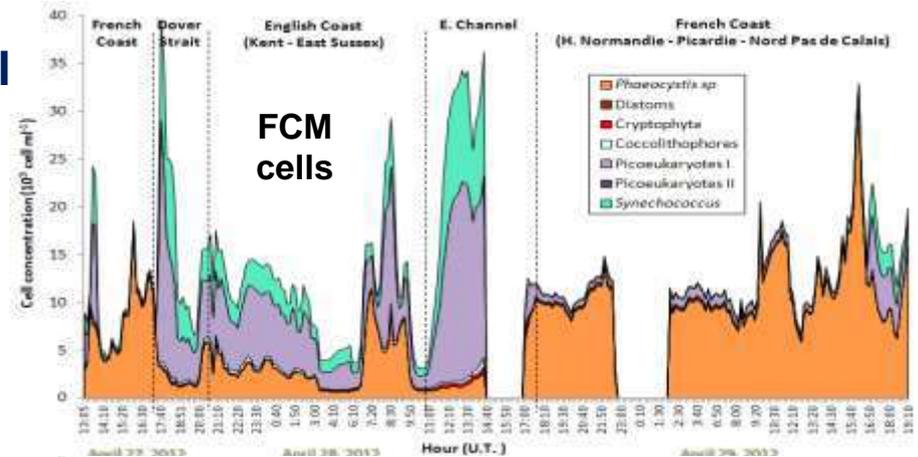
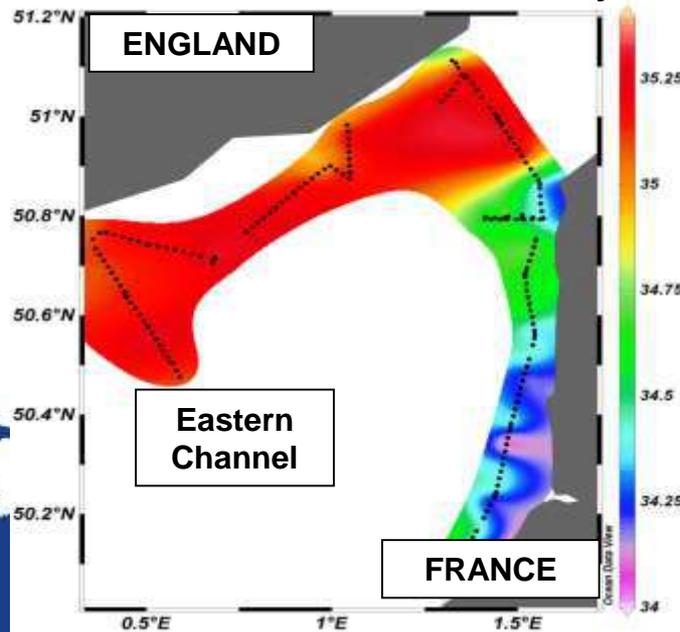
Continuous recording of phytoplankton in eastern Channel coastal waters

DYPHYMA Cruise (Spring 2012)

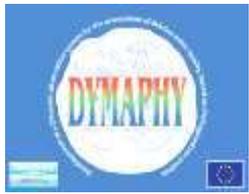


CytoSense (CytoBuoy®)

Salinity



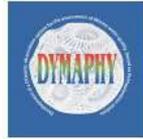
Thèse Doctorat
Simon Bonato
2015, Bonato et al., ECS Science, 2015



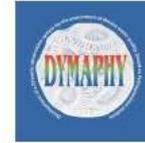
DYMAPHY deliverables



+ PhD Thesis (K. Owen, S. Bonato...) + publications : Bonato et al., 2015; Thyssen et al., 2015, Bonato et al., 2016,



INTERREG IVA 2 Mers Seas Zeeën Crossborder Cooperation Programme 2007-2013



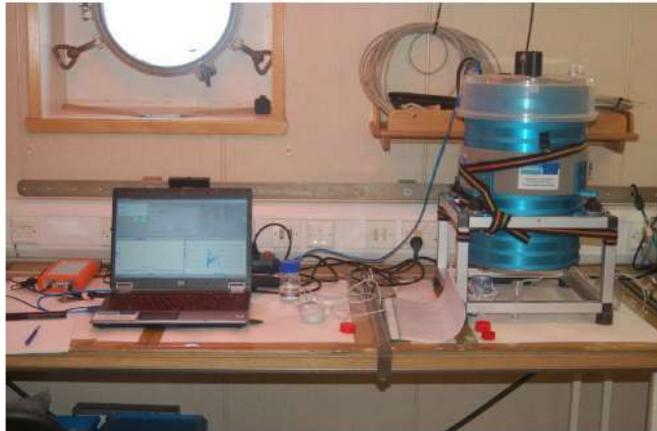
INTERREG IVA 2 Mers Seas Zeeën Crossborder Cooperation Programme 2007-2013
Part-financed by the European Regional Development Fund (ERDF)

DYMAPHY Operational Common Protocol for Pulse-shape recording Flow Cytometry

Recommendations for the validation of
custom built flow cytometers:

Guidance for operators

T.Rutten, M.Rijkeboer, A.Veen, M.Thyssen, S.Bonato, V.Creach, L.F.Artigas



Part-financed by the European Regional Development Fund (ERDF)

Thomas Rutten, Machteld Rijkeboer , Arnold Veen



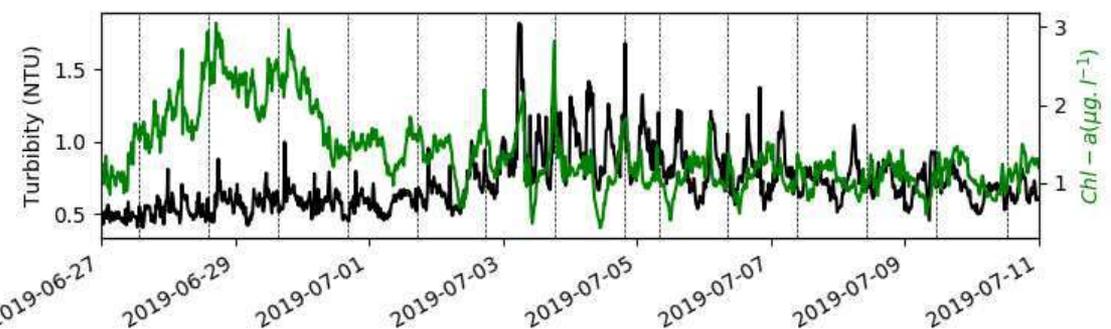
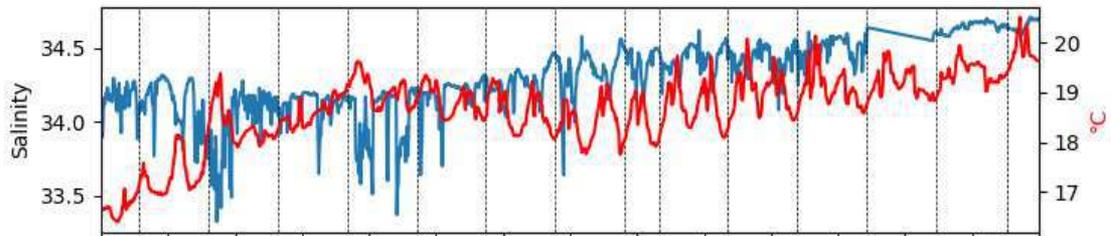
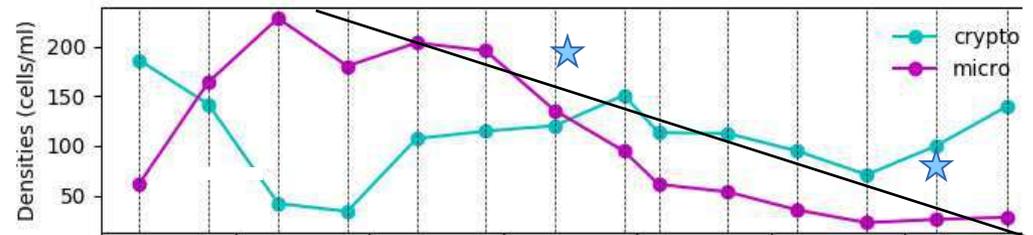
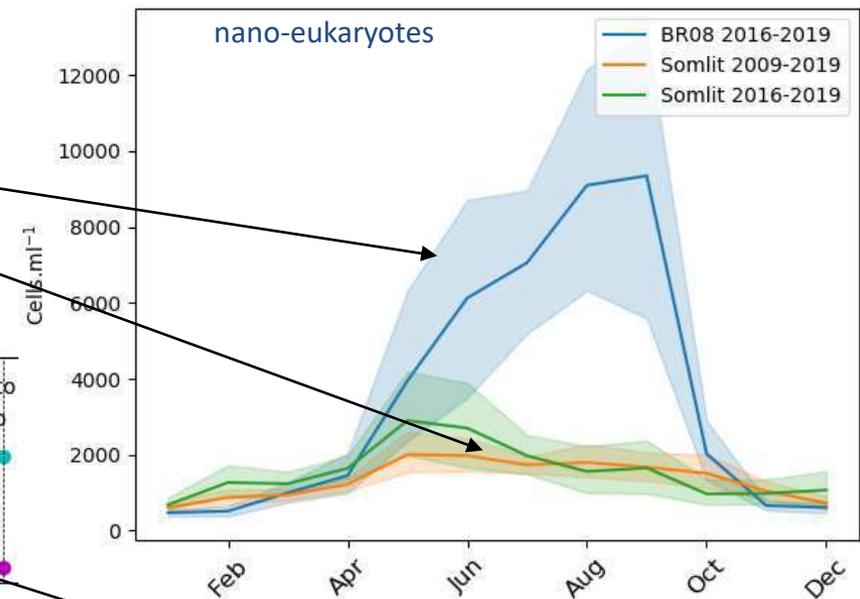
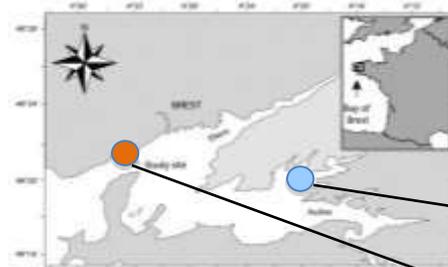
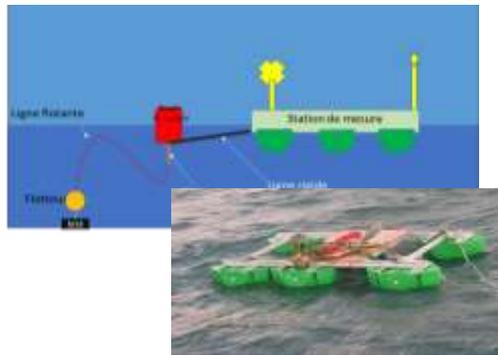
JERICO NEXT deliverables D3.1 & D3.2

+ PhD Thesis (A. Louchart) + publications (Aardema *et al.*, 2019; Marrec *et al.*, 2018; Louchart *et al.*, 2020a, b); De Blok *et al.*, submitted,...

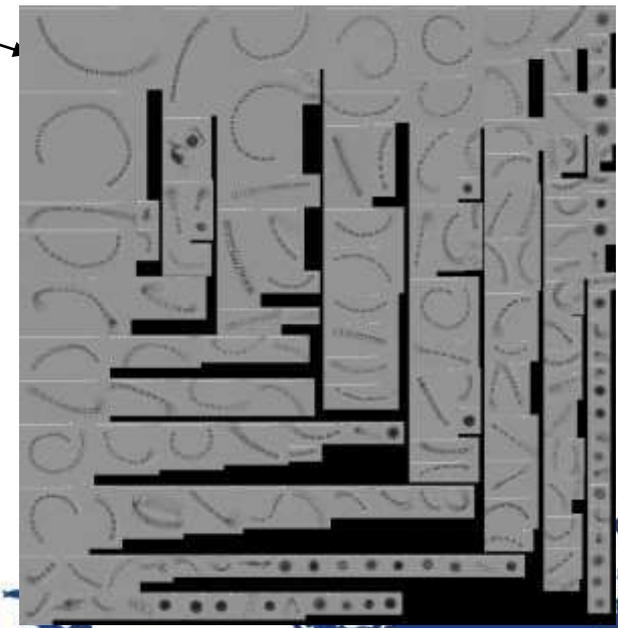


Joint European Research Infrastructure network for Coastal Observatory – Novel European eXpertise for coastal observatoRies - JERICO-NEXT	
Deliverable title	Novel methods for automated <i>in situ</i> observations of phytoplankton diversity
Work Package Title	WP 3
Deliverable number	D3.1
Description	Synthesis report after developments dedicated to the observation of the phytoplankton diversity
Lead beneficiary	SMHI
Lead Authors	Bengt Karlson, Felipe Artigas, Veronique Créach, Arnaud Louchart, Guillaume Wacquet and Jukka Seppälä
Contributors	Hedy Aardema, Michael Brosnahan, Reinhoud de Blok, Pascal Clauvir Florent Colas, Klaas Deneudt, Gérald Grégori, Jacco Kromkamp, Soumay Lahbib, Alain Lefebvre, Fabrice Lizon, , Klas Möller, Emilie Poisson-Caillaud Machteld Rijkeboer, Thomas Rutten, Suvi Rytövuori, Lars Stemmann Melilotus Thyssen, Lennert Tyberghein, and Pasi Ylöstalo.
Submitted by	Bengt Karlson
Revision number	9
Revision Date	4 October 2017
Security	Public

Joint European Research Infrastructure network for Coastal Observatory – Novel European eXpertise for coastal observatoRies - JERICO-NEXT	
Deliverable title	Novel methods for automated <i>in situ</i> observations of phytoplankton diversity and productivity: synthesis of exploration, inter comparisons and improvements
Work Package Title	WP 3
Deliverable number	D3.2
Description	Report on the technical and analytical improvements of innovative techniques and recommendations on their use
Lead beneficiary	CNRS
Lead Authors	Felipe Artigas, Véronique Créach, Emilie Houliez, Bengt Karlson, Fabrice Lizon, Jukka Seppälä, Guillaume Wacquet
Contributors	Hedy Aardema, Michael Brosnahan, Reinhoud de Blok, Pascal Clauvir, Gérald Grégori, Florent Colas, Elisabeth Debusschere, Klaas Deneudt, Jacco Kromkamp, Soumaya Lahbib, Alain Lefebvre, Arnaud Louchart, Klas Möller, Emilie Poisson-Caillaud, Thomas Rutten, Machteld Rijkeboer, Suvi Rytövuori, Lars Stemmann, Melilotus Thyssen, Lennert Tyberghein, Jochen Wollschläger and Pasi Ylöstalo.
Submitted by	Felipe Artigas
Revision number	1



29/06



★ REPHY sampling

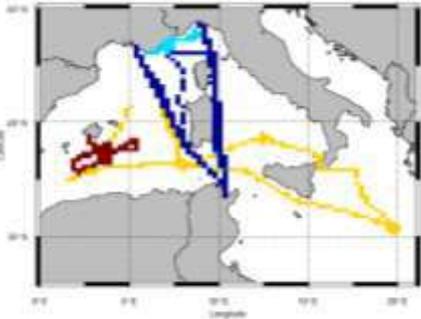
Local time

La cytométrie automatisée au MIO



>2500 échantillons collectés entre 2015 et 2020 lors des campagnes à la mer.

>5500 échantillons de cytométrie en flux collectés entre 2019 et 2020 au SSL@MM, Endoume, Marseille.
(<https://sslamm.mio.osupytheas.fr/>)



OSCAHR
FUMSECK
PEACETIME
CHROME
preBIOSWOT

Responsable: Méliotus Thyssen

Mise en place de la procédure d'interopérabilité dans le cadre de SDC:

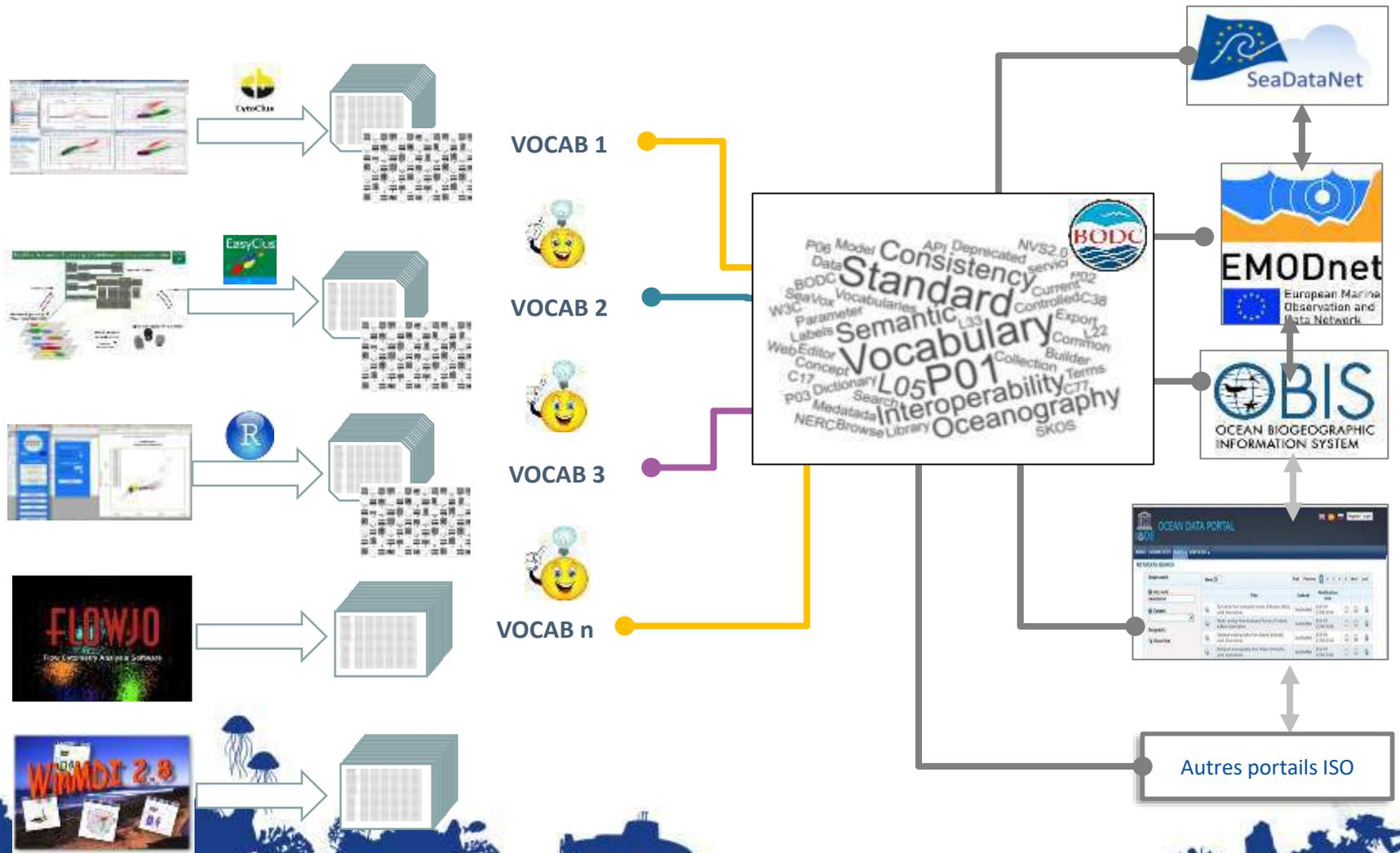
- Adaptation des métadonnées
- vocabulaire commun parmi les experts FCM
- Paramètres optimaux (diffusion, fluorescence)

Création premier vocabulaire interopérable enregistré au BODC

Création d'une base de donnée (CYTOBASE, cf M. Libes)
Application procédures qualité ODV

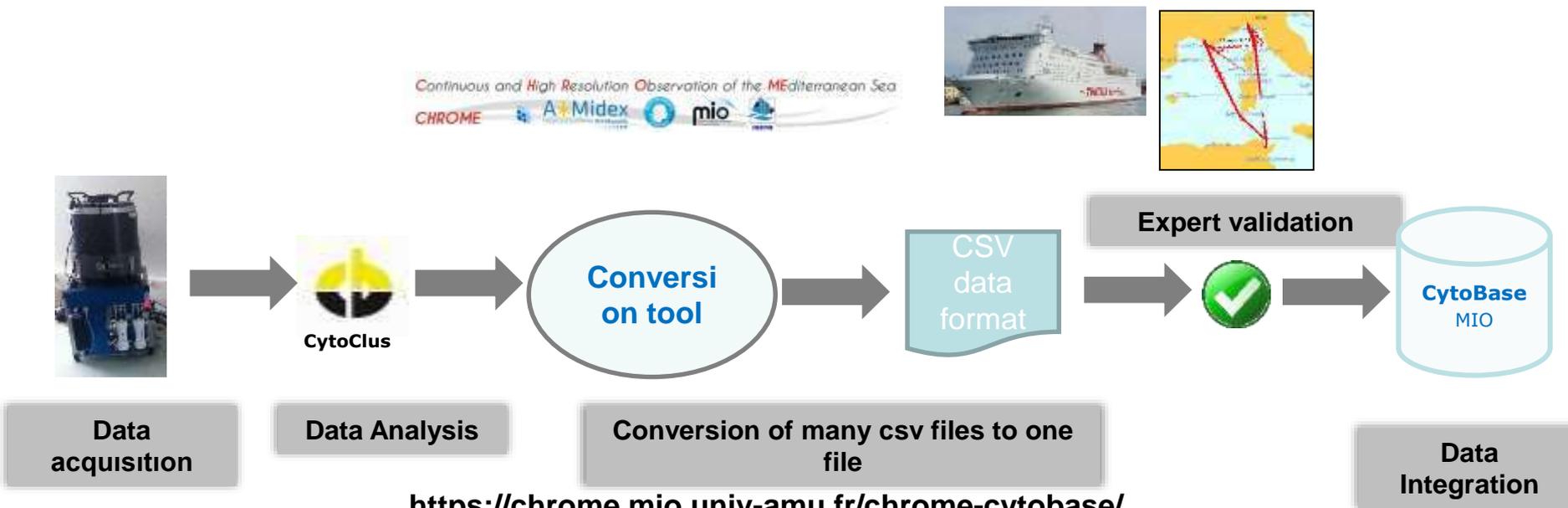


Un vocabulaire interopérable est nécessaire pour permettre aux données issues de la cytométrie en flux d'intégrer des bases de données et être accessibles pour la communauté scientifique



PROJET CHROME 2015-2016 : Soumaya Lahbib + Maurice Libes démarrent la gestion des données cytométrie pour la haute fréquence et créent une Cytobase = couplage métadonnées + données + vocabulaire.

Premier Workflow (cf. M. Libes)



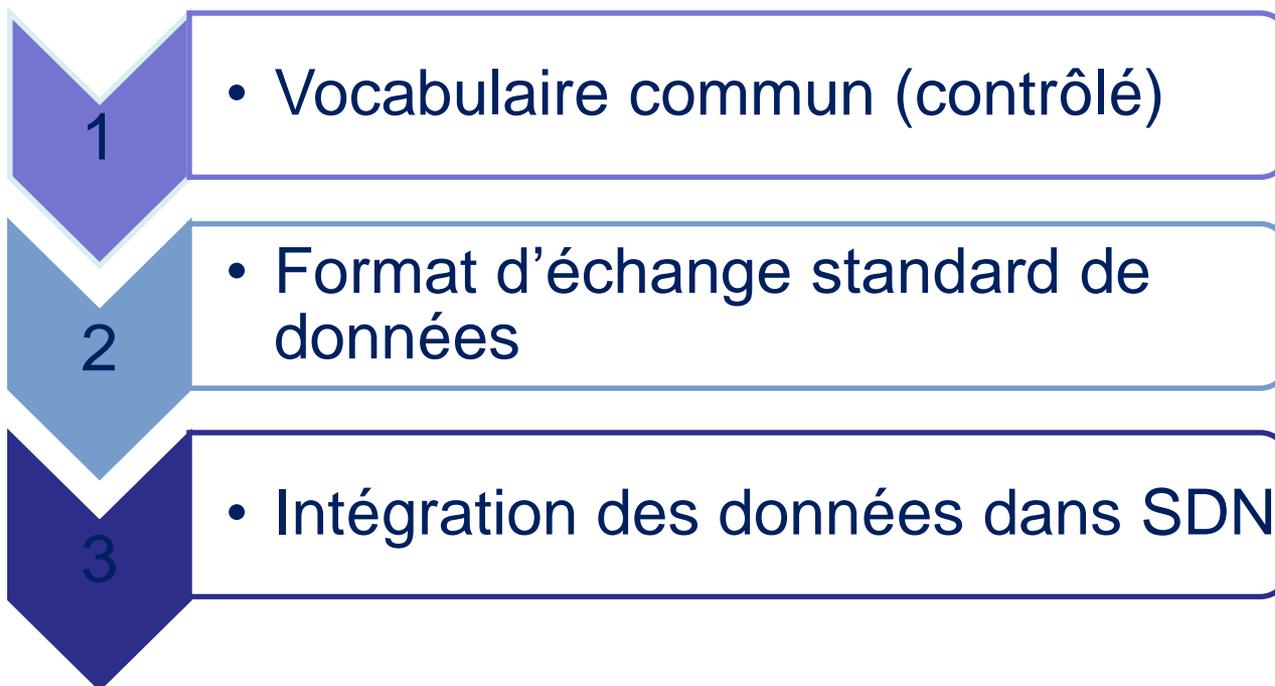
<https://chrome.mio.univ-amu.fr/chrome-cytobase/>
OBJECTIF= RENDRE CES DONNEES ACCESSIBLES A LA COMMUNAUTE SCIENTIFIQUE.

PROJET SEADATACLOUD 2016-2020=accessibilité via les portails de gestion de données marines
Intégration de toutes les données cytométrie-Vocabulaire commun (NERC-BODC)-modèle sémantique (NERC-BODC+EMODNET)-contrôle qualité (ODV)-





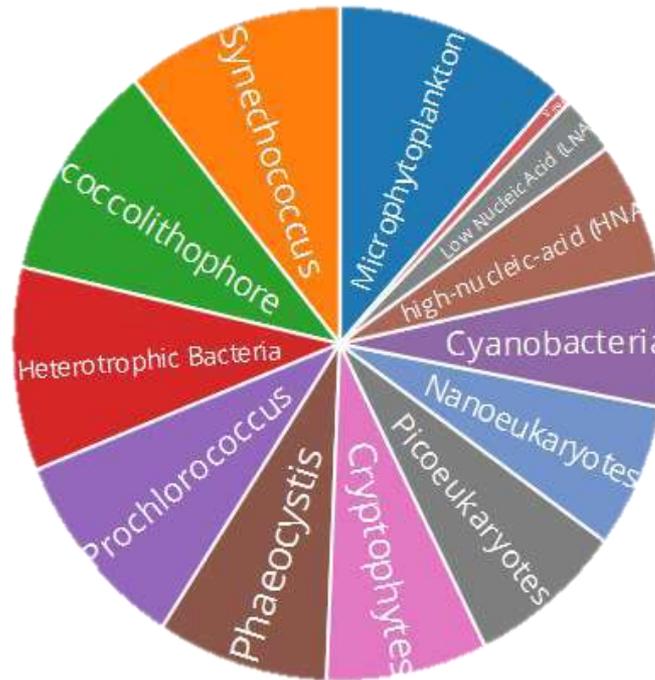
WP9.5.2- Intégration, validation, archivage et accessibilité à long-terme des données de FCM



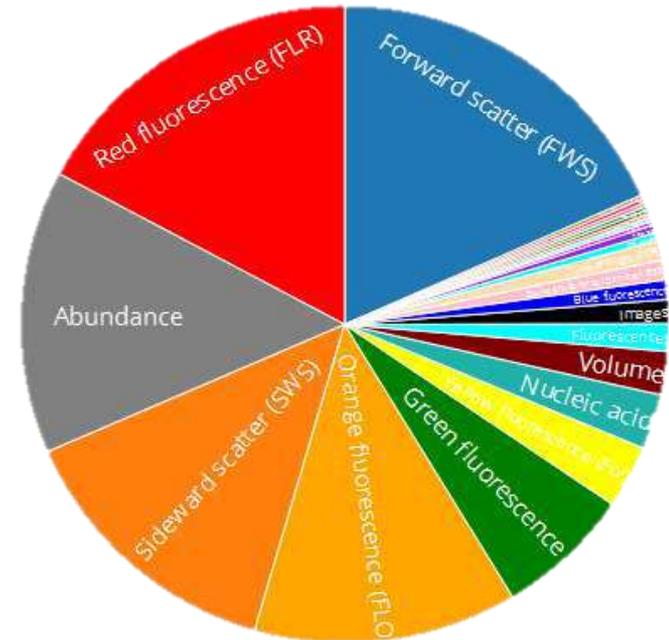
1. Recherche bibliographique de 1983 à 2017



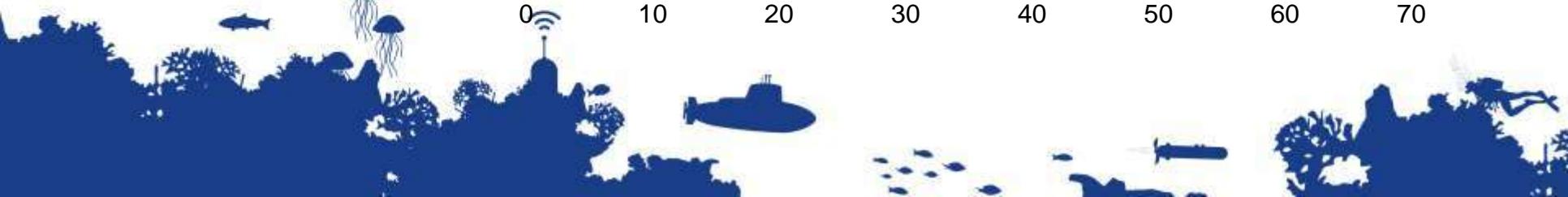
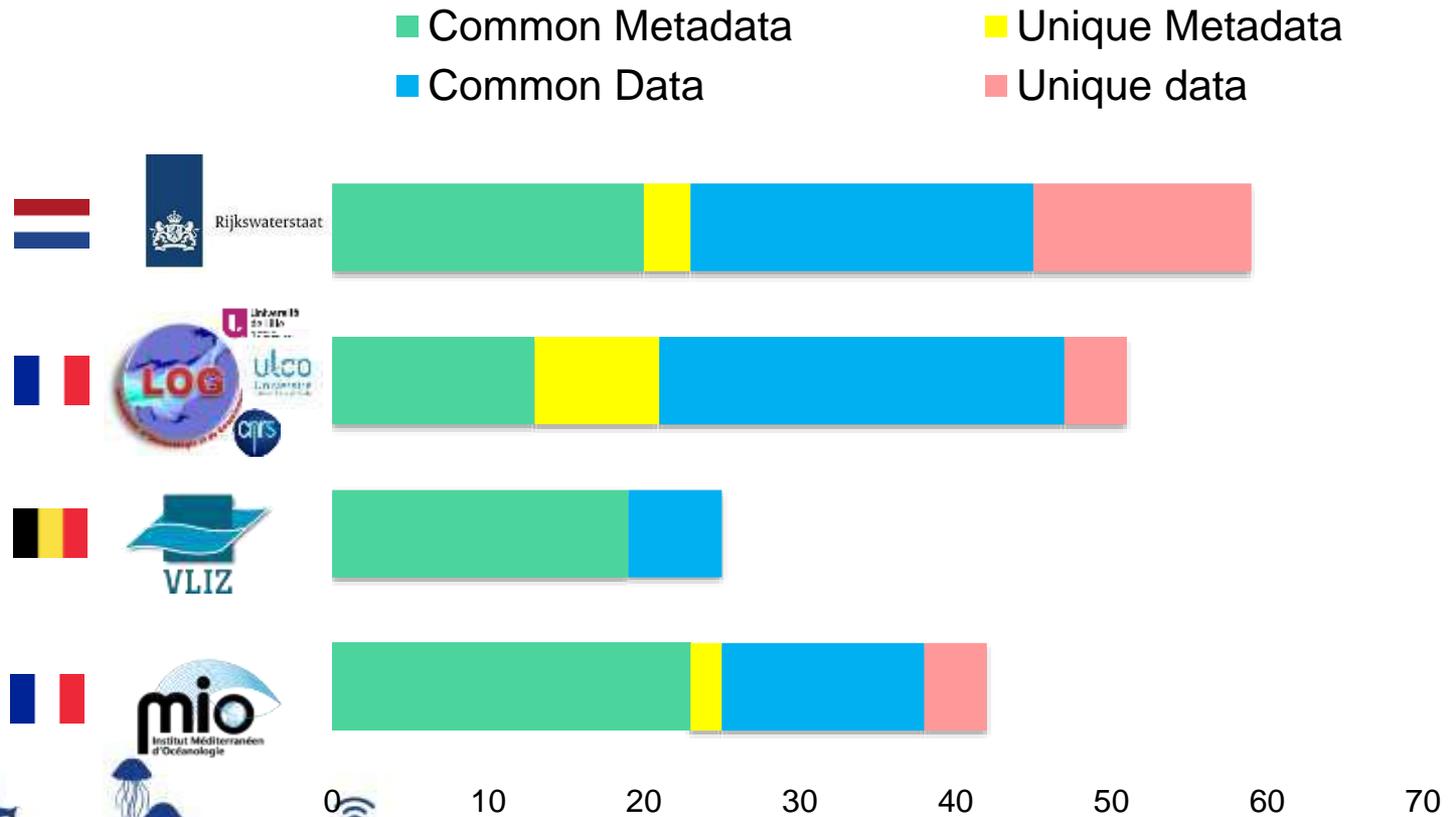
Clusters



Paramètres mesurés



2. Identification des paramètres en commun (exercice au sein du consortium d'utilisateurs du projet JERICO NEXT)

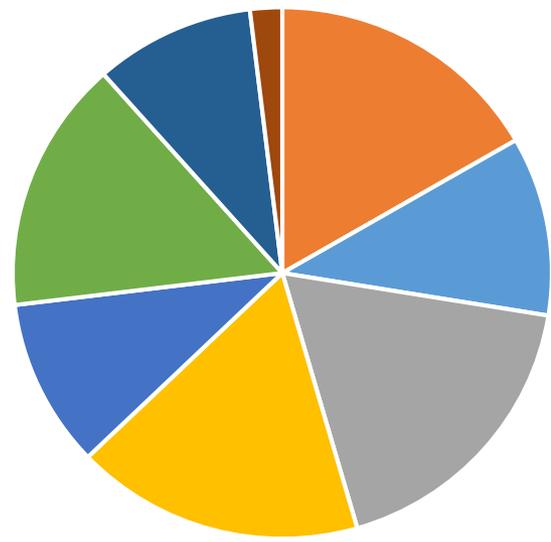


3. Elargissement du consortium

Questionnaire pour les experts : 153 emails - 38 réponses

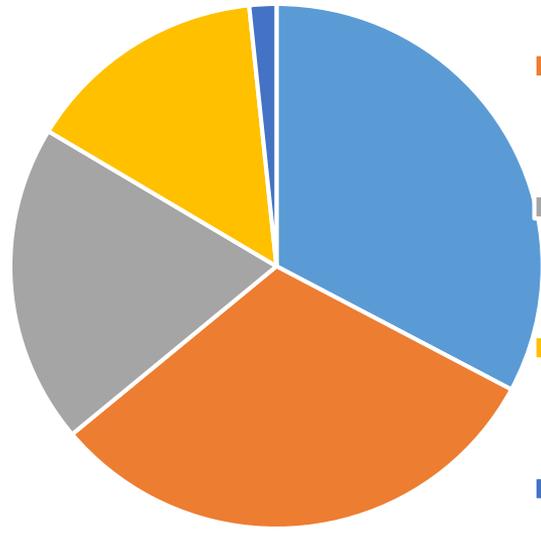
What are the recurrent autotrophic functional groups of your area of study?

- Synechococcus
- Prochlorococcus
- Picoeukaryotes
- Nanoeukaryotes
- Coccolithophore
- Cryptophytes



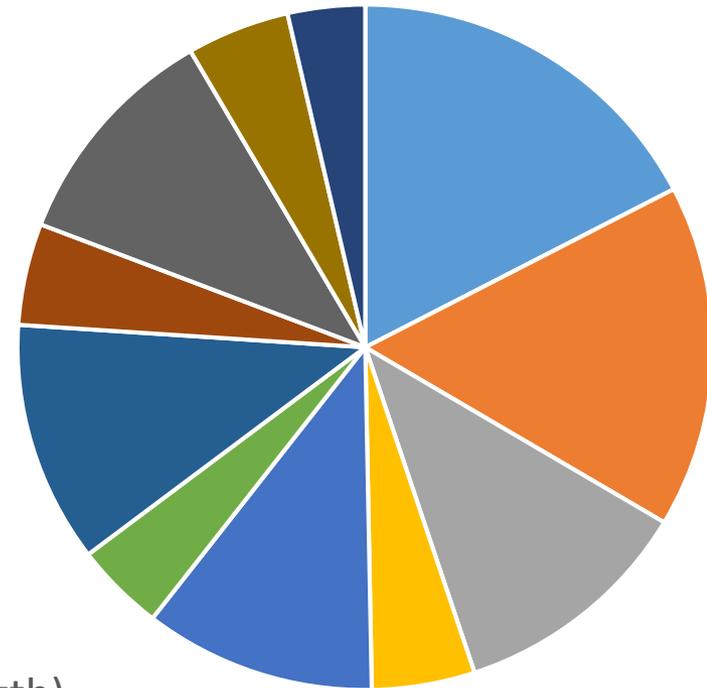
What are the recurrent Heterotrophic groups of your area of study?

- High Nucleic Acid Prokaryotes
- Low Nucleic Acid Prokaryotes
- Nanoflagellates
- Not concerned
- Other (Virus)



What parameters do you export after your clustering?

- Abundance (cell.cm-3)
- Functional group names
- Average Red Fluorescences
- Standard deviation Red Fluorescences
- Average Orange Fluorescences
- Standard deviation Orange Fluorescences
- Average Side Ward Scatter (Area, length)
- Standard deviation Side Ward Scatter (Area, length)
- Average Forward Scatter (Area, length)
- Standard deviation Forward Scatter (Area, length)
- Other



Suite

- Finaliser le vocabulaire (papier en cours)
- ODATIS CES : Mettre en place le **workflow**, les **métadonnées**, gestion des **données campagnes/institutionnelles**
- Intégrer le travail JERICO S3/CES
 - **Best practices** (interopérabilité)
 - Conditions expérimentales appliquées pour pouvoir mettre les données sur le CDS
 - **Quality Control** (qualification des données)
 - Check liste des contrôles effectués pour pouvoir mettre ses données au CDS





Avancement chaîne de traitement de cytométrie en flux

*M. Libes (MIO), M. Sourisseau (IFREMER)
(+ M. Thyssen & G. Gregori (MIO))*

Septembre 2020



Plan

- Etude comparative des workflow de cytométrie existants
 - *MIO/OSU Pytheas et*
 - *Laboratoire d'Ecologie Pélagique Ifremer*
 -
- Réécriture d'une chaîne de traitement de cytométrie complète en Python



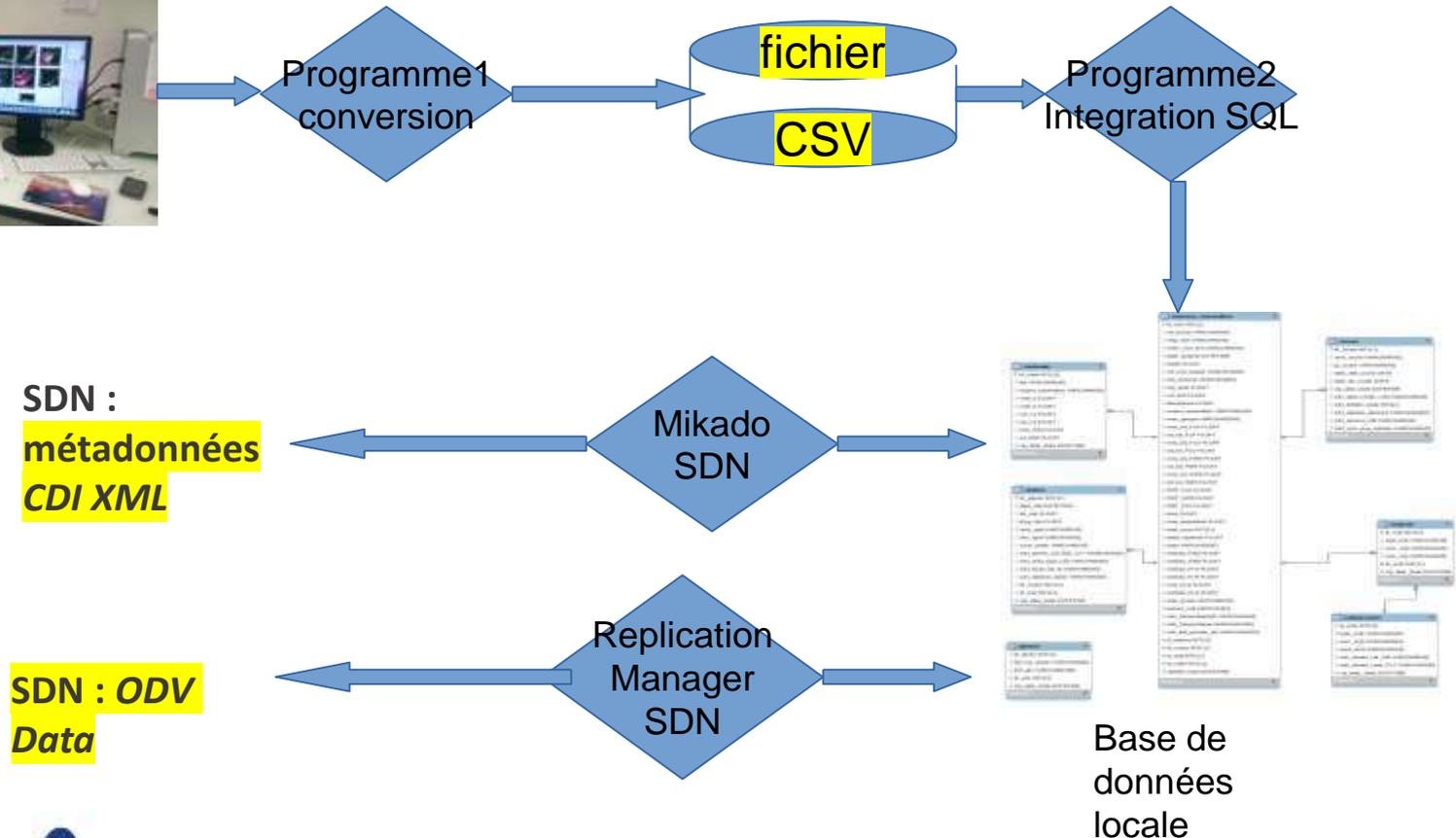
Etude comparative des workflow de cytométrie existants



- **Objectifs**
- Connaître et comparer les chaînes de traitement des données de cytométrie établies au
 - *MIO/OSU Pytheas et*
 - *Laboratoire d'Ecologie Pélagique Ifremer*
- Etablir des recommandations pour profiler une chaîne de traitements commune qui pourrait être recommandée par le CES Odatis et utilisée dans un CDS cytométrie
- Étude détaillée accessible sur :
 - <https://docs.google.com/document/d/1MrRDDRzuA4thz3evZ0kyqsG4RKAjpOFk14EAXiqaM0M/edit?usp=sharing>



Résumé du workflow : *étapes communes importantes*



Conclusions sur fichier CSV

- En sortie des cytomètres il est nécessaire de créer un fichier de données intermédiaire, consolidé ayant fait l'objet d'un contrôle qualité, au format CSV (lignes, colonnes)
- Les entêtes de nos fichiers sont quasiment identiques
- Entête du fichier CSV à l'OSU Pytheas (en rouge les différences avec Laboratoire Ecologie Pélagique.)
- ***Project, Project.starting.Date, Project.ending.Date, PI, Cytometer.ID, Station, Depth, Latitude, Longitude, Study.area, Samples.Operator, Standards.Reference, Clustering.Method, Observation.Type, Platform.Type, Platform.ID, Platform.Nationality, Sampling.Date, Analysis.Date, Standardized.name, Selection.Set, File, Volume, Trigger.Channel, Trigger.level, SWS.amplification, FLO.amplification, FLR.amplification, Abundance, Mean.Total.FWS_varx1, SD.Total.FWS_varx2, Mean.Total.SWS, SD.Total.SWS, Mean.Total.FLR, SD.Total.FLR, Mean.Total.FLO, SD.Total.FLO, Mean.Max.FLR, SD.Max.FLR, Beta.0, Beta.1, Mean.Length, SD.Length***



Conclusions fichier CSV

- Différences
 - ordre des colonnes :
 - l'ordre des colonnes diffère : peu important d'un point de vue programmatique,
 - l'important est le libellé des colonnes : quasiment identique
Cependant pour une homogénéité totale on pourrait envisager d'avoir aussi le même ordre de colonnes
 - intitulés des colonnes :
 - Selection.Set (aka Expert_Nomination) absent dans le fichier CSV Ifremer
 - Mean.Max.FLR, SD.Max.FLR, parametres absents dans le fichier CSV Ifremer

•



Conclusions fichier CSV

Au laboratoire d'Ecologie pelagique

- Le fichier CSV créé comporte déjà la somme des abondances des groupes standardisés. Il correspond à la “vue” MIKADO_<le_nom_dela_campagne>. de la BD Cytobase de l'OSU Pytheas

Au MIO/OSU : on a un niveau de détail supplémentaire avec un fichier intermédiaire avant la somme des abondances des groupes:

- le 1er fichier CSV en sortie du cytomètre comporte une colonne “Selection Set” correspondant aux détails des groupes déterminés par l'expert cytométriste dans les analyses d'une date de prélèvement donnée
- Le 2eme fichier CSV fournit pour une date de prélèvement donnée “sampling_date”,
 - i) la somme des abondances des groupes expertisés par le cytométriste pour une date de prélèvement donnée
 - ii) Les moyennes pondérées des propriétés optiques de ces groupes pour une date de prélèvement donnée

Conclusions BD

- Les fichiers CSV produits sont intégrés dans une BD locale (optionnel)
- Parmi les différences on note la présence d'un peu plus de métadonnées dans les tables de la BD Cytobase OSU Pytheas, dont certaines utilisant le vocabulaire contrôlé du BODC pour l'intégration dans SeadataNet
- Peu importe le schéma relationnel de la BD des laboratoires. Ce qui importe ce sont :
 - les fichiers terminaux produits en sortie
 - Format interopérable ODV ou NetCDF demandé
 - Les métadonnées nécessaires
 - L'utilisation du vocabulaire contrôlé du BODC

Le vocabulaire interopérable pour la cytométrie est en cours de publication et issue d'un consensus entre 32 experts internationaux.



Conclusions : vocabulaire contrôlé

- Dans le cadre d'une nécessaire interopérabilité, il est important d'utiliser des tables du vocabulaire standard existantes dans SDN:
- C17_SDN_Embarcation_code.csv : nom et code des bateaux : 35LU "Le Suroit"
- L06_SDN_Embarcation_Type.csv : type du bateau : research vessel
- L22_SDN_cytometre_code.csv : le code des cytometres : TOOL1209, CytoBuoy CytoSense flow cytometer,
- et surtout la table des codes des groupes de cytométrie
 - F02_SDN_Cytometrie.csv : code des groupes de phytoplancton F0200003, Synechococcus,



Conclusions



- Quelques différences mineures dans le fichier CSV issu des Cytomètres :
- Dans le cadre d'un CDS :
 - Aller vers une standardisation des fichiers CSV issus des analyses de données cytométriques. Se mettre d'accords sur le format interne de ces fichiers et sur les métadonnées nécessaires (le format CSV "Mikado" étant déjà une bonne base initiale)..
 - Aller dans le sens de l'utilisation de vocabulaire contrôlé standardisé issus d'un consensus d'experts
 - Le format national de sortie de cytométrie en flux repose déjà sur le travail effectué dans le cadre de SDC, et est effectué en parallèle par un consortium européen dans le cadre du projet JERICOS3. Une méthode nationale n'a de valeur que si elle est internationale, sinon, il faut tout recommencer
- Quelques différences dans les BD :
 - Le modèle de données interne des BD locales est peu important pour autant que les processus de traitement logiciels fournisse les fichiers de sortie corrects avec les métadonnées attenantes.
 - Dans le cadre d'un CDS : fournir des fichiers de cytométrie complets au format ODV ou NetCDF pour Odatis
- Dans le cadre d'un CES Odatis, se mettre d'accord sur le format et contenu des fichiers de cytométrie interopérables produits (quels paramètres doivent être rendus publics?)
- Fournir des outils logiciels en Python aux laboratoires pour compiler leurs données de cytomètres de manière identiques : avoir une batterie d'outils logiciel s'adaptant aux différentes générations de cytomètres

Travail réalisé au MIO/OSU Pytheas

WP Chaîne de traitement des données de cytométrie

- Mars-Juillet 2020 : Réécriture en *python* de toute la chaîne de traitement avec un stagiaire (*L. Izard, M. Libes*) pour des les programmes *Cytoclus3* et *Cytoclus4*
- On traite les données brutes des modèles « *cytoclus4* »
- La chaîne effectue désormais un grand nombre de contrôles qualité sur les données brutes issues des cytomètres
- Intégration des campagnes *chrome*, *peacetime*, *fumseck*



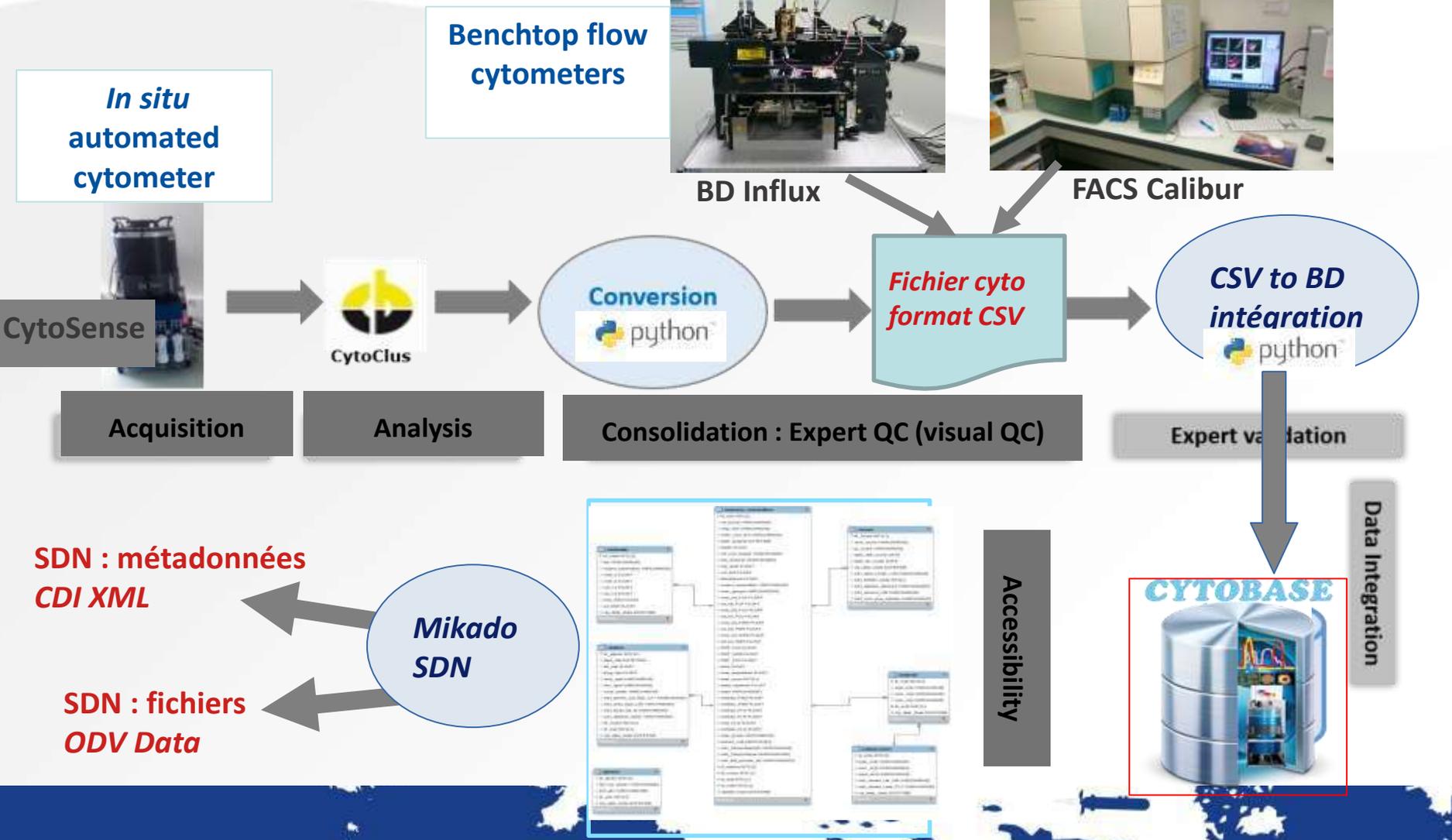


SeaDataNet

PAN-EUROPEAN INFRASTRUCTURE FOR OCEAN & MARINE DATA MANAGEMENT



FCM Data management Workflow



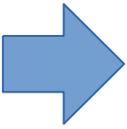
Travaux réalisés et futurs à DYNECO/PEL

Chaîne de traitement des données de cytométrie

Juillet 2020 – Mars 2021

- Intégration des images associées à l'imagerie en flux dans la BD locale (fini)
- Intégration des données de 2 suivis temporels (Octobre)
- Mise en place de procédure d'annotation (Novembre).
- Mise en place de procédure d'export de ces images vers des plateformes d'imagerie (Janvier).





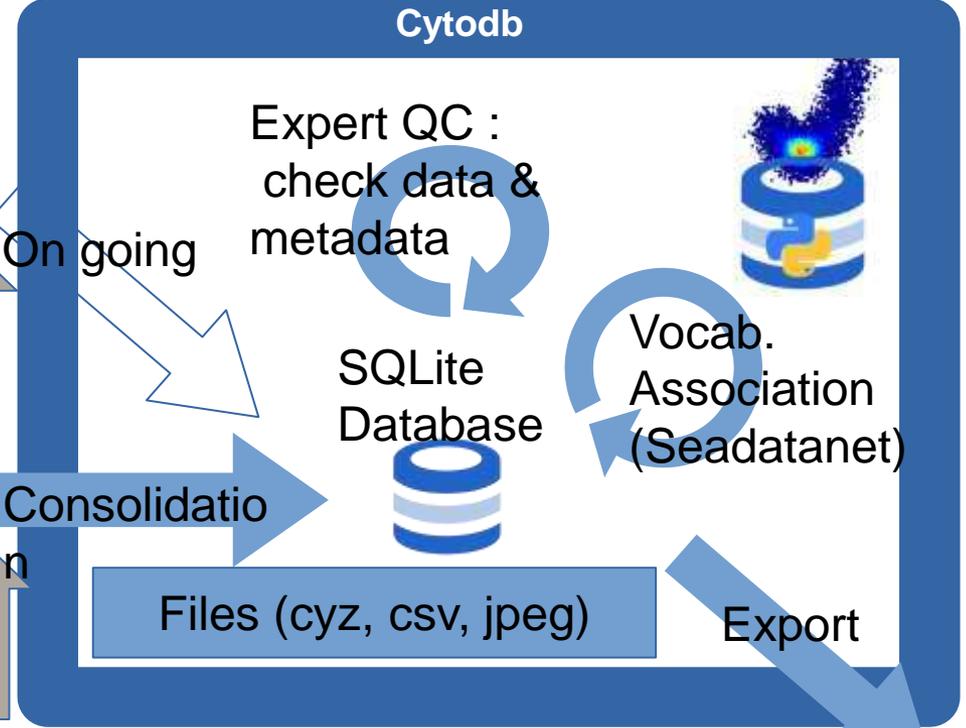
Data acquisition (Lab or autonomous)

Manual Analysis

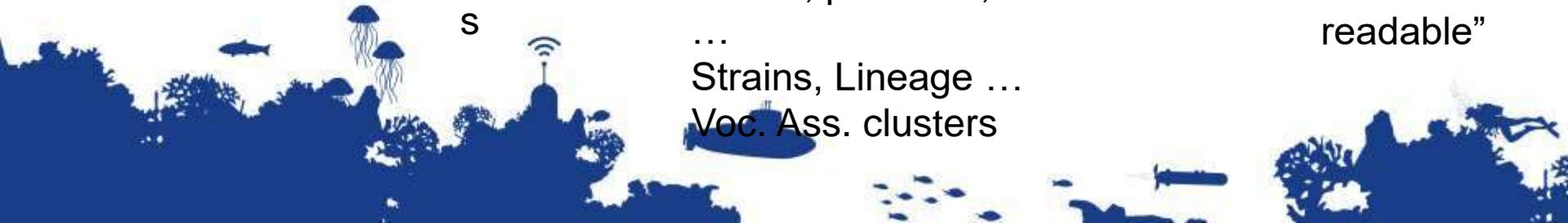
Inputs :
Cruise, positions, ...
Strains, Lineage ...
Voc. Ass. clusters

Export

Export



CSV "Mikado readable"





Merci de votre attention !

