

PROPOSITIONS ET ELEMENTS DE REFLEXION POUR LA CREATION D'UNE BANQUE D'ECHANTILLONS ET D'EMPREINTES D'ANALYSE CHIMIQUE NON CIBLEE (NTS)

Sophie Lardy-Fontan, Anne Togola

Avec la contribution de Valéria Dulio

Août 2019

Document final



Avec le soutien de :

**AGENCE FRANÇAISE
POUR LA BIODIVERSITÉ**

ÉTABLISSEMENT PUBLIC DE L'ÉTAT



Contexte de programmation et de réalisation

Ce rapport a été réalisé dans le cadre du programme RDI/RSP lot C pour l'année 2018-2019, au titre de l'activité 7 du réseau de surveillance prospective « Evaluation de techniques permettant l'archivage physique ou virtuel d'informations chimiques sur les milieux aquatiques, et l'analyse rétrospective des échantillons ».

Auteur (s) :

Sophie Lardy-Fontan
LNE
sophie.lardy-fontan@lne.fr

Anne Togola
BRGM
a.togola@lne.fr

Approbateur:
Sophie Vaslin-Reimann
LNE
sophie.vaslin-reimann@lne.fr

Vérification du document :

Marina Coquery
Irstea
marina.coquery@irstea.fr

Les correspondants

AFB : Pierre-François STAUB, pierre-francois.staub@afbiodiversite.fr
Olivier PERCEVAL, olivier.perceval@afbiodiversite.fr

LNE : Sophie Vaslin-Reimann, sophie.vaslin-reimann@lne.fr

Référence du document : Sophie Lardy et al. PROPOSITIONS ET ELEMENTS DE REFLEXION POUR LA CREATION D'UNE BANQUE D'ECHANTILLONS ET D'EMPREINTES D'ANALYSE CHIMIQUE NON CIBLEE (NTS) 54 pages

Droits d'usage :	<i>Accès libre</i>
Couverture géographique :	<i>International</i>
Niveau géographique :	<i>National</i>
Niveau de lecture :	<i>Professionnels, experts</i>
Nature de la ressource :	<i>Document</i>

GLOSSAIRE.....	8
1. INTRODUCTION / CONTEXTE	9
1.1 Généralités	9
1.2 Note de position : De l'intérêt des banques d'échantillons pour la surveillance des milieux aquatiques	9
1.3 Note de position : De l'intérêt du screening non ciblé dans le Réseau de Surveillance Prospective	10
1.4 Contexte et objectif du travail	11
2. METHODOLOGIE.....	12
3. CONCEPT, ENJEUX ET RECOMMANDATIONS	13
3.1 organisation de l'infrastructure.....	13
3.2 Coordination / gestion de l'infrastructure	13
3.3 Valorisation de l'infrastructure	14
3.4 Accessibilité	14
3.5 Propriété : droits d'auteurs et droits de propriété.....	15
3.6 Promotion.....	16
3.7 Interopérabilité	17
3.8 informatisation et Traçabilité dans les banques	17
4. SYNTHESE DES RECOMMANDATIONS.....	18

Liste des annexes :

Annexe 1:Support de présentation de la réunion du 07 janvier 2019	21
Annexe 2: La gestion des données : contexte et exigences.....	54

PROPOSITIONS ET ELEMENTS DE REFLEXION POUR LA CREATION D'UNE BANQUE D'ECHANTILLONS ET D'EMPREINTES D'ANALYSE CHIMIQUE NON CIBLEE (NTS)
S. LARDY-FONTAN, A. TOGOLA

RESUME

Cette action vient alimenter l'activité 7 du réseau de surveillance prospective « **Evaluation de techniques permettant l'archivage physique ou virtuel d'informations chimiques sur les milieux aquatiques, et l'analyse rétrospective des échantillons** ». Il complète la **Note de positionnement en appui au Réseau de Surveillance Prospective– Rapport AQUAREF 2017– 16 p ,1 annexe.**

Ce travail établit des recommandations pour soutenir le bon fonctionnement de futures infrastructures Banques d'échantillons environnementaux de manière générale. Il aborde également certains points particuliers liés aux banques d'empreintes NTS « spectrothèque ». Le document met en avant les enjeux liés à la traçabilité de l'information et surtout la capitalisation des données /informations nouvellement acquises à partir des échantillons physiques ou empreintes NTS intégrés dans ces infrastructures notamment au sein des BDD de surveillance existantes.

Le rapport propose des réflexions et des recommandations sur les points suivants/

- Organisation de l'infrastructure
- Coordination et gestion de l'infrastructure
- Valorisation de l'infrastructure
- Accessibilité des échantillons et empreintes NTS stockés dans l'infrastructure
- Propriété : droits d'auteurs et droits de propriété de l'infrastructure et des échantillons empreintes NTS qui y sont stockés
- Promotion de l'infrastructure
- Interopérabilité avec les autres banques d'échantillons et les BDD de surveillance
- Informatisation et Traçabilité des échantillons et des données au sein de l'infrastructure et au-delà.

Ces éléments de réflexions et recommandations ont été repris afin de proposer des modes de gestion et de fonctionnement d'infrastructures de types banques d'échantillons et spectrothèque.

Mots clés (thématique et géographique) :

Banques d'échantillons, spectrothèque, organisation, traçabilité

PROPOSALS AND ELEMENTS OF REASONING FOR THE CREATION OF AN ENVIRONMENTAL SPECIMEN BANK AND NON TARGET SCREENING (NTS) FINGERPRINTS BANK
S. LARDY-FONTAN, A. TOGOLA

ABSTRACT

This action feeds activity 7 of the prospective monitoring network "Evaluation of techniques allowing the physical or virtual archiving of chemical information on aquatic environments, and the retrospective analysis of samples". It complements the Positioning paper "Note in support of the Prospective Surveillance Network" - AQUAREF Report 2017- 16 p, 1 appendix.

This work establishes recommendations to support the smooth operation of future environmental specimen banks and infrastructures in general. It also discusses some specific points related to NTS fingerprinting banks. The document highlights the issues related to the traceability of information and especially the capitalization of newly acquired data / information from physical samples or NTS fingerprints embedded in these infrastructures, particularly within existing surveillance databases.

The report proposes reasoning and recommendations on the following points :

- Organisation of the infrastructure
- Coordination and management of the infrastructure
- Enhancement of the infrastructure
- Accessibility of samples and NTS fingerprints stored in the infrastructure
- Ownership: copyright and property rights of the infrastructure and NTS fingerprint samples stored there
- Promotion
- Interoperability with other ESB and monitoring BDDs
- Computerization and traceability of samples and data within the infrastructure and beyond.

These elements of reasoning and recommendations have been taken up in order to propose modes of management and functioning for a national ESB and spectrothèque.

Key words (thematic and geographical area):
ESB, spectrothèque, traceability, organization

GLOSSAIRE

Banque d'échantillons : infrastructure permettant d'archiver des échantillons (ex. sols, sédiment, biote, échantillonneurs intégratifs passifs, ADN) et des empreintes NTS à très long terme (plusieurs années)

Banque de données : ensemble de données élémentaires issues d'un ou plusieurs dispositifs de collecte organisés par traitement informatique. Elle permet d'extraire des données et de les mettre à disposition des utilisateurs. Les données élémentaires des dispositifs de collecte peuvent être situées dans plusieurs banques de données. Dans certains cas la banque de données peut contenir des données élaborées. Source : d'après AFB et OIEau

DMP : Data Management Platform = plateforme de gestion des données en français

Donnée brute : donnée élémentaire issue d'une mesure (index de compteur, hauteur d'eau au droit d'une station de jaugeage, etc.) qui n'a encore été ni validée, ni organisée dans une banque de données, ni interprétée. Source : d'après Ministère chargé de l'environnement

Donnée traitée NTS : donnée issue du retraitement d'une empreinte NTS

Donnée ouverte : se dit d'une démarche qui vise à rendre réellement ré-utilisables les données produites par un organisme, en autorisant leur réutilisation libre et gratuite, et en les proposant selon des formats facilitant leur réutilisation. Source : d'après AFB

Echantillothèque : banque d'échantillons physiques ex. sols, sédiment, biote, échantillonneurs intégratifs passifs, ADN

Empreinte NTS (ou échantillons virtuel): fichier de données brutes acquis dans des conditions analytiques spécifiques par une méthode d'analyse non ciblée..

NTS : Non target screening - analyse non ciblée

PGD : plan de gestion de données (DMP) est un document formalisé explicitant les modalités de production, de traitement, de protection et de diffusion des données, la propriété intellectuelle, les choix réalisés en termes d'archivage, ainsi que les coûts associés à la gestion des données dans le cadre de la recherche sur projet

Retraitement : extraction à partir d'un logiciel des informations présentes dans l'empreinte NTS selon les objectifs visés et qui aboutit à des données traitées NTS

Spectrothèque : banque d'empreintes NTS

1. INTRODUCTION / CONTEXTE

1.1 GENERALITES

Cette action vient alimenter l'activité 7 du réseau de surveillance prospective « **Evaluation de techniques permettant l'archivage physique ou virtuel d'informations chimiques sur les milieux aquatiques, et l'analyse rétrospective des échantillons** ».

En 2017, une note stratégique regroupant le positionnement d'AQUAREF ¹.sur l'intérêt d'une banque d'échantillons physiques dans le cadre de la surveillance prospective et sur les perspectives d'utilisation des techniques de screening non ciblé dans le cadre d'études prospectives et pour la priorisation des substances émergentes a été rédigée. Les éléments d'intérêt mis en exergue dans le cadre de ce travail sont repris ci-dessous. ²

1.2 NOTE DE POSITION : DE L'INTERET DES BANQUES D'ECHANTILLONS POUR LA SURVEILLANCE DES MILIEUX AQUATIQUES

Contexte

Les banques d'échantillons ont pour principe commun d'archiver des échantillons à très long terme (plusieurs années). Dans le contexte de la transition écologique, ce type d'infrastructures est reconnu comme un outil puissant d'évaluation des politiques publiques de gestion/remédiation de l'état de l'environnement notamment au sens DCE pour les milieux aquatiques. La plupart des Etats qui mènent des politiques environnementales ambitieuses se sont dotées de telles infrastructures : Allemagne, Suède, Norvège, en Europe ; Etats-Unis, Japon, hors Europe. En France, la mytilothèque gérée par l'Ifremer ou l'écothèque développée par l'ANDRA en sont des exemples notables. Au-delà de leur première raison d'être, les banques d'échantillons servent aussi aux desseins de la recherche, en renforçant les plateformes pluridisciplinaires d'excellence.

Les échantillons environnementaux archivés dans ces banques sont aussi divers que les objectifs pour lesquels ils sont utilisés et les pays d'où ils viennent. Enfin, le design d'une banque d'échantillons – choix des matrices, échantillonnage, préparation et conditions de conservation - doit répondre aux objectifs de sa mise en oeuvre. De plus, son exploitation nécessite de définir des règles éthiques et des droits/conditions d'accès.

L'exploitabilité des banques d'échantillons repose sur des processus d'assurance qualité et de métrologie performants et maîtrisés afin d'optimiser les protocoles de stockage mais également la préparation et l'analyse initiale des échantillons, pour garantir à long terme la comparabilité des mesures et d'en tirer des conclusions fiables.

Historiquement, dans le domaine environnemental, les banques d'échantillons étaient constituées d'échantillons physiques biotiques ou abiotiques solides, pertinents pour la surveillance chimique. Si ces bases, opérationnelles depuis près de 50 ans, demeurent tout à fait pertinentes, elles seront complétées dans un futur proche par de nouvelles approches qu'il reste à construire et consolider. On pourra plus particulièrement citer :

- La conservation d'échantillons intégratifs passifs autorisant la conservation indirecte de matrices jusqu'ici non conservées, notamment l'eau ;
- L'émergence de nouvelles techniques analytiques : approche de screening non ciblé (NTS) associée aux progrès informatiques et internet autorisant de nouvelles stratégies de conservation dites virtuelles ou dématérialisées. Celles-ci ne visent pas, en l'état

¹ Togola Anne et al – Note de positionnement en appui au Réseau de Surveillance Prospective– Rapport AQUAREF 2017– 16 p ,1 annexe.

² Togola Anne et al – Note de positionnement en appui au Réseau de Surveillance Prospective– Rapport AQUAREF 2017– 16 p ,1 annexe.

actuel des connaissances, à se substituer à la conservation d'échantillons physiques mais bien à la compléter.

- L'émergence des approches « omiques » (génomique, protéomique, métabolomique) susceptibles de mettre en évidence et identifier de nouveaux biomarqueurs (d'exposition, d'effet ou de susceptibilité), générer de nouvelles connaissances sur le plan mécanistique (modes d'action), ou encore élaborer de nouveaux outils de toxicologie prédictive pour aider à l'identification des dangers.

La note présente par ordre croissant de difficultés, les différentes propositions et axes de réflexion.

De l'intérêt des banques d'échantillons dans le cadre de la surveillance de l'état chimique

Une banque d'échantillons nationale pourrait permettre de :

- ✓ Soutenir les besoins des suivis en tendances
- ✓ Contribuer à l'identification des sources de pollution
- ✓ Contribuer à évaluer l'impact de rupture technologique sur la qualité et l'interprétation des données

De l'intérêt des banques d'échantillons dans le cadre du Réseau de Surveillance Prospective- surveillance chimique

Une banque d'échantillons nationale pourrait permettre de :

- ✓ Mettre en œuvre des méthodes de mesures mûres à posteriori de l'échantillonnage
- ✓ Revenir sur des données historiques a posteriori pour aiguiller et soutenir les futurs exercices de priorisation (itération)
- ✓ Permettre d'acquérir des chroniques sur le long terme et d'établir des tendances.

De l'intérêt des banques d'échantillons pour la Surveillance de la biodiversité et de l'intégrité biologique des milieux

Surveillance de l'intégrité biologique des milieux

Les banques d'échantillons sont également de puissants outils de recherche pour les approches biomarqueurs/bioindicateurs pour :

- ✓ Permettre de mettre en œuvre des méthodes de mesures mûres a posteriori
- ✓ Comprendre la toxicocinétique / organotropisme de certains contaminants
- ✓ Contribuer à évaluer l'impact de ruptures technologiques sur la qualité et l'interprétation des données
- ✓ Revenir sur des données historiques a posteriori pour soutenir les futures évolutions réglementaires notamment

Surveillance de la biodiversité

La génomique environnementale est un domaine de recherche émergent qui regroupe l'ensemble des connaissances acquises sur les organismes et les écosystèmes présents et passés, par l'analyse de la séquence et de l'expression des gènes, génomes et métagénomes. Analysant les potentialités et les capacités d'adaptation des organismes vivants, la génomique environnementale permet de déterminer leur biologie, leurs traits fonctionnels et leurs interactions avec l'environnement dans ses dimensions biotique et abiotique. Un nombre croissant de biologistes utilise l'ADN pour la détection des espèces et la biosurveillance, contournant ou atténuant ainsi le besoin d'échantillonner des organismes vivants.

La conservation pour :

- ✓ Permettre un archivage de la biodiversité ex situ
- ✓ Revenir sur des données historiques a posteriori
- ✓ Permettre de mettre en œuvre des méthodes de mesures mûres a posteriori

1.3 NOTE DE POSITION : DE L'INTERET DU SCREENING NON CIBLE DANS LE RESEAU DE SURVEILLANCE PROSPECTIVE

L'utilisation de la spectrométrie de masse pour l'analyse non ciblée a connu un essor considérable ces dernières années dans de nombreux domaines, dont celui de

l'environnement. Un premier rapport AQUAREF, produit en 2015, propose un état de l'art et des préconisations pour l'inclusion de ce nouvel outil dans la surveillance des milieux aquatiques. La présente note reprend certains de ces éléments et fait un focus sur leur intérêt dans le cadre du Réseau de Surveillance Prospective.

Une bancarisation virtuelle de l'échantillon

Les traitements a posteriori des analyses non ciblées permettront d'étudier des évolutions temporelles et/ou spatiales, ou de rechercher de nouvelles molécules. Pour cela, l'information brute doit être stockée de manière pérenne et associée aux métadonnées liées à l'échantillonnage, l'analyse, les contrôles-qualité etc...

La bancarisation de ces analyses est donc essentielle et doit être construite (au-delà des possibilités apportées par les constructeurs) et harmonisée. La faisabilité et l'intérêt d'une centralisation nationale des données afin de créer une véritable banque virtuelle d'échantillons doivent être étudiés.

Des approches comparables sont actuellement en construction au niveau européen, notamment dans le cadre du réseau NORMAN. Les interactions entre une action nationale et cette action européenne doivent être approfondies afin d'améliorer la mise en commun d'outils de bancarisation et de faciliter l'échanges de données.

1.4 CONTEXTE ET OBJECTIF DU TRAVAIL

Ce travail entrant dans le champ du RDI "Réseau de Surveillance Prospective, RSP", il est restreint à des propositions en lien avec les besoins d'une infrastructure de banque d'échantillons (échantillons physiques et spectrothèque) au service de la surveillance de la contamination chimique de l'environnement, notamment sur les points suivants : organisation, accessibilité, promotion, valorisation.

Même si ce travail n'a pas vocation à aborder la question de la bancarisation au sens « échanges de données et Base de données de la surveillance », certains éléments de réflexion seront introduits dans ce travail.

Au-delà des aspects techniques, méthodologiques et documentaires qui seront traités dans certaines actions du RSP, l'exploitabilité d'une banque d'échantillons implique de réfléchir aux questions de propriétés des échantillons et des données associées, aux questions éthiques, ainsi qu'aux questions d'accessibilité et de promotion. Il est absolument essentiel d'aborder ces aspects dès la conception de la banque d'échantillons car ils conditionnent la manière dont cette dernière sera organisée. Ce rapport présente des propositions pour l'établissement de chartes éthiques et de règles de fonctionnement de futures infrastructures de banques d'échantillons environnementaux. Ce travail aborde également certains points particuliers liés aux banques d'empreintes NTS.

A ce jour, il n'y a aucune décision ou stratégie définie au niveau national que ce soit par le MTES ou l'AFB au sujet d'infrastructures de type échantillothèque ou spectrothèque (Les rencontres de l'Onema n°29³) :

- création d'une infrastructure nouvelle ou fédération d'infrastructures existantes en France ;
- quels services pour quels missions/enjeux :
 - infrastructure pour l'archivage de l'environnement/milieux (sauvegarde patrimonial)
 - infrastructure en soutien de la recherche/innovation

³ <https://professionnels.afbiodiversite.fr/sites/default/files/Rencontres29.pdf>

- infrastructure pour démontrer l'impact des politiques publiques sur la qualité de l'environnement/milieux;...
- etc.

De ce fait, les points spécifiques traités dans ce document et les propositions associées devront être précisées et/ou complétées, le cas échéant, selon les décisions futures.

Pour les points plus détaillés sur la construction de la spectrothèque il faut se référer au livrable « Scenarii pour la mise en place d'une spectrothèque : banque d'empreintes NTS »⁴

2. METHODOLOGIE

Une recherche sur les sites internet ou par le biais d'échanges écrits a été réalisée auprès des principales infrastructures ou réseaux nationaux, européens et extra européens qui hébergent / coordonnent une banque d'échantillon (communication orale par Koschorreck_ICESB2015 ; Les rencontres de l'Onema n°29⁵) avec pour objectif de recueillir des informations/documents sur les points suivants : propriétés des échantillons et des données associées, questions éthiques, questions d'accessibilité et de promotion. Les principaux réseaux ou infrastructures qui ont été consultés ou sollicités sont les suivants :

En France :

- ANDRA : Ecothèque (France)
- IFREMER : Mytilothèque (France)
- INRA : Conservatoire des sols Pédothèque (France)
- INRA : ADNthèque Genols (France)
- Muséum national d'Histoire Naturelle MNHN (France)

A l'interface France - Europe

- EBMRC France Ressources biologiques marines (France, réseau Européen)
- Observatoire de Recherche sur la Qualité de l'Environnement du Grand Sud-Ouest Européen (France, réseau Européen)

A l'international :

- NIST- Marine Environmental Specimen Bank (Etats-Unis)
- es-BANK, Center for Marine Environmental Studies, Ehime University (Japon)
- NIVA: The Environmental Specimen Bank (ESB Norway) (Norvège).
- Swedish Museum of Natural History (Suède), etc.
- ESB Bilbao (Pays Basque espagnol)

Pour les points particuliers liés à la banque d'échantillons virtuels, des échanges avec Nikiforos Alygizakis et Jaroslav Slobodnik qui travaillent à la construction de la Digital Freeze Sample Platform (DFSP) dans le cadre des activités de NORMAN, ont été organisés.

L'ensemble de ces éléments a été présenté lors d'une réunion BRGM, INERIS, LNE, AFB, DEB qui s'est déroulée le 7 janvier 2019 et sont disponibles dans l'annexe 1.

⁴ Togola et al., Scenarii pour la mise en place d'une spectrothèque : banque d'empreintes NTS 2019, .

⁵ <https://professionnels.afbiodiversite.fr/sites/default/files/Rencontres29.pdf>

3. CONCEPT, ENJEUX ET RECOMMANDATIONS

3.1 ORGANISATION DE L'INFRASTRUCTURE

En règle générale, deux types d'organisation sont observés. Soit la banque d'échantillons est adossée à un organisme (exemple : agence environnementale, instituts publics tel que l'IFREMER ou l'INRA, etc.), soit l'infrastructure est intégrée comme un outil d'une suprastructure agréant notamment des plateformes techniques et technologiques (exemples : EMBRC, ESB Bilbao, etc.). Dans ce deuxième cas, l'infrastructure est positionnée comme un outil en soutien à la recherche et l'innovation favorisant ainsi le développement des procédés pour répondre aux attentes de la communauté scientifique (meilleure conservation des échantillons, systèmes de traçabilité innovants, implémentation de rupture technologique, etc.).

Ce choix de configuration est critique et en lien avec la stratégie nationale sur les infrastructures de type échantillothèque ou spectrothèque, en suspens. Il conditionne également les modes de fonctionnement et d'organisation qui seront décrits dans les paragraphes suivants.

3.2 COORDINATION / GESTION DE L'INFRASTRUCTURE

Comme toute infrastructure, le besoin impérieux de disposer d'instances stratégiques, décisionnelles et de pilotage multipartites est évident.

Pour la mise en place au niveau national de futures banques d'échantillons physiques ou virtuels, il est recommandé de mettre en place 2 instances :

Un comité d'orientation stratégique :

Cette instance multipartite pourrait être composée de représentants :

- des ministères (MTES mais éventuellement Ministère de la recherche, Ministère de la santé)
- de l'AFB,
- d'infrastructures préexistantes (Ecothèque, Pédothèque, etc.), si le choix de la mise en réseau d'infrastructures était fait,
- et inclure un comité scientifique international constitué d'experts.

Ce comité aurait, parmi ses missions, la détermination des orientations stratégiques et des ressources nécessaires à sa mise en œuvre, l'animation du réseau et des équipes, l'ouverture internationale (mise en réseau) et la définition de la politique d'accessibilité et de propriété intellectuelle (DMP) (cf. paragraphe 3.4).

Un comité de pilotage

Cette instance multipartite pourrait être constituée des administrateurs/représentants de l'infrastructure, d'opérationnels et de scientifiques et serait en charge de la bonne mise en œuvre des décisions stratégiques définies par le comité d'orientation stratégique.

Si plusieurs infrastructures étaient mises en réseau, les comités de pilotage déjà existant au sein de chacune d'entre elles seraient à représenter ; parmi les principales missions de cette instance devrait être inclus : recevabilité/inclusion des échantillons dans la banque d'échantillons, l'évaluation des demandes d'accès aux échantillons bancarisés, l'analyse des résultats générés et in fine inclusion dans une banque de données de référence, les questions d'accessibilité (3.4) et de gestion dans le temps des spécimens et des données stockés dans l'infrastructure (3.4), et de promotion (3.6).

3.3 VALORISATION DE L'INFRASTRUCTURE

Plus qu'un simple système de fourniture d'échantillons, une infrastructure de banque d'échantillons physiques et/ou empreintes NTS est une archive qui agrège des données dont la valeur est difficilement quantifiable. En effet,

1. La valeur est subjective : elle dépend de l'intérêt porté par l'utilisateur, directement ou après transformation ;
2. La valeur est co-construite : elle augmente à partir de l'instant où elle rentre dans un processus de transformation, notamment à travers des études croisées, capables d'exprimer des concepts, d'où l'importance de la collaboration et de la coordination dans le processus de la valorisation des données ;
3. La valeur est potentielle : elle donnera ou pas un avantage futur à ceux qui la détiennent.

C'est pour cette raison qu'il est important que les connaissances acquises au travers de l'exploitation de l'infrastructure y reviennent pour l'enrichir et permettre une acquisition croisée de connaissances (notion de data mining précédemment évoquée...). Ces points devront être précisés dans la politique d'accessibilité (3.4).

3.4 ACCESSIBILITE

La question de l'accessibilité englobe un spectre assez large de points spécifiques dont certains ont déjà été abordés ou seront abordés dans les différents paragraphes de ce document.

Le rôle du comité stratégique et du comité de pilotage sont centraux dans ce processus.

Les points suivants devront être discutés et définis au sein de ces deux instances :

- Organisation des demandes : demande toute l'année « au fil de l'eau » ou appel annuel.
- Droits d'accès :
 - graduels selon le niveau d'implication dans la surveillance en France.
 - implication possible (ou obligatoire) de membres du comité de pilotage dans les projets visant à l'exploitation des échantillons physiques/empreintes NTS de l'infrastructure.
- Définition du coût d'accès aux échantillons, participation aux coûts d'envoi pour les échantillons physiques.
- Définition de la politique d'accessibilité (présentation PGD) et notamment du point de retour des d'informations /connaissances produites par l'exploitation des archives vers l'infrastructure
- Définition du mode de citation de l'infrastructure : citation dans la valorisation des résultats/informations

Pour les demandeurs, il faudra mettre en place des formulaires de demande d'échantillons. Ces documents devront notamment inclure les éléments suivants :

- Description du demandeur (réfèrent technique /équipe projet ; niveau d'expertise) et du bénéficiaire final.
- Motivation de la demande et notamment justification précise du besoin d'accès à des archives (échantillons physiques ou empreintes NTS) et pas uniquement à de simples supports d'études (plus-value de l'échantillon archivé par rapport à un nouvel échantillon).
- Motivation des quantités demandées pour les échantillons physiques.
- Signature de la DMP.

L'organisation mise en œuvre au niveau du NIST à laquelle sont adossées les banques d'échantillons environnementaux aux Etats-Unis illustre un mode d'organisation structuré répondant aux enjeux et questions abordés au travers de ce document : définition de niveaux d'accès différents, pilotage graduée, articulation de la base, etc. (Figure 1).

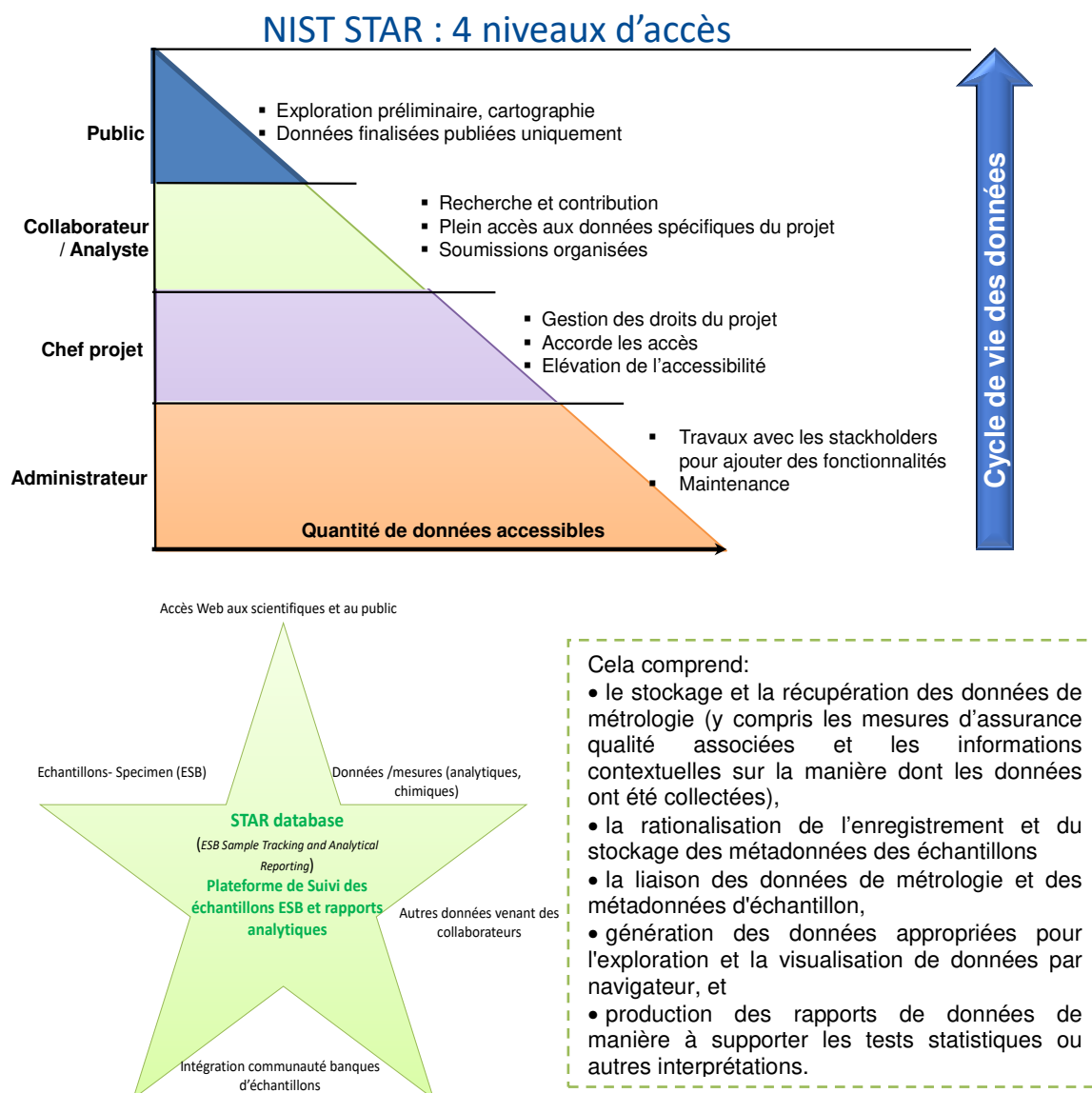


Figure 1 : Banque d'échantillons Etats-Unis, NIST : Organisation

3.5 PROPRIETE : DROITS D'AUTEURS ET DROITS DE PROPRIETE

En Europe, les bases de données possèdent leur propre protection juridique, depuis la Directive européenne du 11 mars 1996.

La première protection concerne uniquement la forme de la base, son architecture, et est conditionnée comme pour toute autre œuvre par une condition d'originalité. La base doit avoir un choix d'indexage original pour être protégée par le droit d'auteur.

La deuxième protection, spécifique aux bases de données, concerne la matière contenue par la base. Le droit sui generis est rangé dans la catégorie des droits voisins du droit d'auteur, droit de propriété incorporelle ad hoc, donnant des prérogatives patrimoniales au producteur de la base. Mais comme pour le droit d'auteur, l'exercice du droit est attaché à une condition. Ici, il ne s'agit pas d'originalité, mais de valeur économique : la base doit

avoir été l'objet d'un investissement qualitativement ou quantitativement substantiel. Le producteur de la base de données peut donc interdire à tout utilisateur l'extraction d'éléments quantitativement ou qualitativement substantiels de la base, ou l'extraction systématique de celle-ci. À noter qu'il est indifférent que la base soit publique ou non. **Les données publiques restent publiques et sont libres de droit**; mais **ce qui est protégé est leur assemblage en un schéma particulier**, selon l'idée que le tout vaut plus que la somme des composants.

Concernant les banques d'échantillons physiques, la question est peu complexe. En effet, elle repose sur le processus d'acquisition de spécimens (objets physiques/biologiques) en vue de leur archivage et, le cas échéant, des mesures ou empreintes NTS qui pourront être réalisées/ acquises sur ces dits échantillons.

Concernant les banques d'empreintes NTS, la question est plus complexe. En effet, cette approche pose la question de **la définition de ce qu'est la donnée**.

Une donnée peut être définie comme ce qui est connu et qui sert de point de départ à un raisonnement ayant pour objet la détermination d'une solution à un problème en relation avec cette donnée. Une empreinte NTS (dans le cas du NTS équivalent à un spectrogramme) est une donnée non interprétée émanant d'une source primaire (dans le cas du NTS équivalent à un couplage HRMS), ayant des caractéristiques liées à celle-ci (dans le cas du NTS équivalent aux spécifications opératoires de la méthode d'acquisition mise en œuvre par le couplage HRMS) et qui n'ont été soumises à aucun traitement ou toute autre manipulation (dans le cas du NTS: soustraction de blancs, etc.).

Il conviendra donc d'intégrer, dans les Appels à projets ou cahiers des charges impliquant l'utilisation des banques, des plans de gestion des données (PGD) qui constituent un élément clé d'une bonne gestion des données. Un PGD décrit le cycle de vie de gestion des données pour celles qui doivent être collectées, traitées et/ou générées par un projet. Afin de rendre les données de recherche trouvables, accessibles, interopérables et réutilisables (FAIR), un PGD devrait comprendre des informations sûres :

- le traitement des données de la recherche pendant et après la fin du projet ;
- les données qui seront collectées, traitées et/ou générées ;
- la méthodologie et les standards qui seront appliqués ;
- la possibilité ou non de partage/libre accès ;
- la manière d'assurer la curation et la conservation des données (y compris après la fin du projet).

Le point particulier des empreintes NTS devra être inclus. Ce point est précisément décrit dans le livrable ad'hoc⁶.

3.6 PROMOTION

La question de la promotion est assez hétérogène au sein des infrastructures et notamment leur rôle/services. Certaines infrastructures qui sont au service de la communauté scientifique sans à priori disposent de catalogues d'échantillons en ligne afin de permettre une valorisation des ressources disponibles et ainsi faciliter la demande d'accès. Au contraire, d'autres infrastructures qui ont une stratégie plus restrictive d'accès ont choisi de ne pas communiquer de manière exhaustive sur les spécimens archivés. Selon les décisions qui seront prises et notamment si la décision de mettre en réseau des infrastructures existantes est retenue, **l'interopérabilité sera un point critique de même**

⁶ Togola et al., Scenarii pour la mise en place d'une spectrothèque : banque d'empreintes NTS 2019, .

que la mise en place de format d'archivage harmonisé permettant de garantir une traçabilité univoque des spécimens et empreintes NTS.

3.7 INTEROPERABILITE

Un outil permettant de **générer des catalogues d'échantillons accessibles en ligne** s'inscrit dans le besoin d'interopérabilité des banques de données avec pour objectifs :

- d'améliorer les échanges de données autour des banques d'échantillons afin d'optimiser la valorisation des ressources.
- De faciliter l'intégration de sources de données hétérogènes dans un système de recherche et de navigation uniforme.

Cela répond également aux questions de promotion/visibilité et d'accessibilité aux échantillons constituant cette base.

3.8 INFORMATISATION ET TRAÇABILITE DANS LES BANQUES

L'exploitabilité et la valeur de toutes banques d'échantillons (échantillons physiques ou empreintes NTS) repose sur la qualité et fiabilité des échantillons physiques, des empreintes NTS et des données associées qui la constituent. **Au-delà de la simple traçabilité strictement liée à l'identification/codification de tout échantillon / empreinte NTS, l'enjeu de la connaissance et de la maîtrise de l'ensemble des actions/processus qui ont conduit à la production de cet échantillon / empreinte NTS est critique de même que les données contextuelles indispensables à l'exploitation future des données nouvellement générées.**

Echantillothèque (banque d'échantillons physiques) :

- Programme de surveillance / étude / programme de recherche, finalité de l'échantillonnage, stations, géo référencement,
- Echantillonnage / prélèvement : les champs/éléments qui sont déjà associés aux données de surveillance semblent suffisants
- Analyse : les champs/éléments qui sont déjà associés aux données de surveillance semblent suffisant
- Données/mesures réalisées dont analyses non ciblées (empreinte NTS) doivent être bancarisées pour garantir la connaissance de l'ensemble des données contextuelles (incluant les QA/QC) et propres à chaque échantillon.

Les formats d'échange de type « EDILABO » SANDRE qui sont utilisés dans le cadre de la surveillance réglementaire sont une réponse à ces besoins et devraient permettre également de garantir l'interopérabilité au niveau français. Cependant ils ne sont, à l'heure actuelle, pas utilisés par la communauté de recherche. Cela impliquerait la transformation de données pour les rendre compatibles avec ces formats d'échanges, ce qui ne semble pas insurmontable si l'ensemble des éléments (champs) nécessaires sont pris en compte. La difficulté résiderait alors dans l'interopérabilité avec des infrastructures hors France. En Europe le format wise pourrait être une alternative.

Spectrothèque (banque d'empreintes NTS) :

Même si ce travail n'a pas vocation à aborder la question de la bancarisation au sens échanges de données et Base de données de la surveillance, il est utile de préciser que dans le contexte des échanges de données EDILABO :

Pour les analyses non ciblées, un travail a été engagé avec le SANDRE et les parties prenantes afin de pouvoir tracer l'information qu'une analyse de ce type a été réalisé sur un échantillon X. Cette information sera qualitative.

Ainsi, il a été décidé de créer trois paramètres SANDRE de nature physique qualitatifs **Empreinte chimique de TYPE A,B,C,D selon le principe du couplage HRMS**

utilisé (LC-ESI/neg-HRMS, LC-ESI/pos-HRMS, GC/CI/HRMS, GC/APCI/HRMS). Ces paramètres qualitatifs prennent pour valeurs possibles au niveau du résultat d'analyse :

- «0» : réalisée
- «1» : non réalisée

Le code remarque de l'analyse prendra pour valeur «1» par défaut, par convention, même si le code remarque «0» (Analyse non faite) aurait pu être utilisé. Les champs obligatoires existants permettront de savoir où les données brutes sont stockées.

En outre, concernant les données spécifiques aux empreintes NTS, certaines personnes sollicitées ont suggéré **d'associer un DOI (Digital Object Identifier) ou un url à ces données brutes**. Le DOI consiste en l'attribution d'identifiants pérennes aux objets issus de la recherche est proposé : il répond à la fois à la question de la propriété et de la traçabilité. **Cette possibilité devra également être considérée.**

Pour les échantillons physiques, une **réflexion/approche comparable pourrait être conduite**.

Il conviendra également de définir les prérequis qui conditionneront la recevabilité de l'échantillon / empreinte NTS pour intégrer l'Echantillothèque ou la spectrothèque.

4. SYNTHESE DES RECOMMANDATIONS

Ce travail établit des recommandations pour soutenir le bon fonctionnement de futures infrastructures Banques d'échantillons environnementaux de manière générale. Il aborde également certains points particuliers liés aux banques d'empreintes NTS « spectrothèque ». Le document met en avant les enjeux liés à la traçabilité de l'information et surtout la capitalisation des données / informations nouvellement acquises à partir des échantillons physiques ou empreintes NTS intégrés dans ces infrastructures notamment au sein des BDD de surveillance existantes.

Le rapport propose des réflexions et des recommandations sur les points suivants :

- Organisation de l'infrastructure
- Coordination et gestion de l'infrastructure
- Valorisation de l'infrastructure
- Accessibilité des échantillons et empreintes NTS stockés dans l'infrastructure
- Propriété : droits d'auteurs et droits de propriété de l'infrastructure et des échantillons empreintes NTS qui y sont stockés
- Promotion de l'infrastructure
- Interopérabilité avec les autres banques d'échantillons et les BDD de surveillance
- Informatisation et traçabilité des échantillons et des données au sein de l'infrastructure et au-delà.

Ces éléments de réflexions et recommandations ont abouti à la proposition de modes de gestion qui sont résumés dans les figures suivantes. La

Figure 2 présente une proposition de gestion possible pour la partie banque d'échantillons. La

Exploitation des empreintes NTS de la spectrothèque

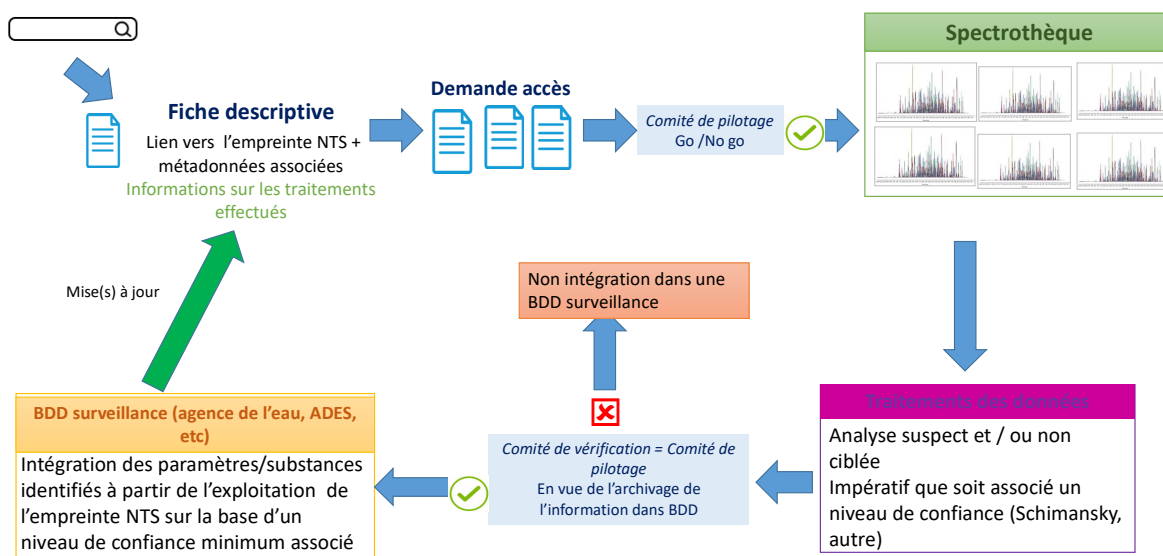
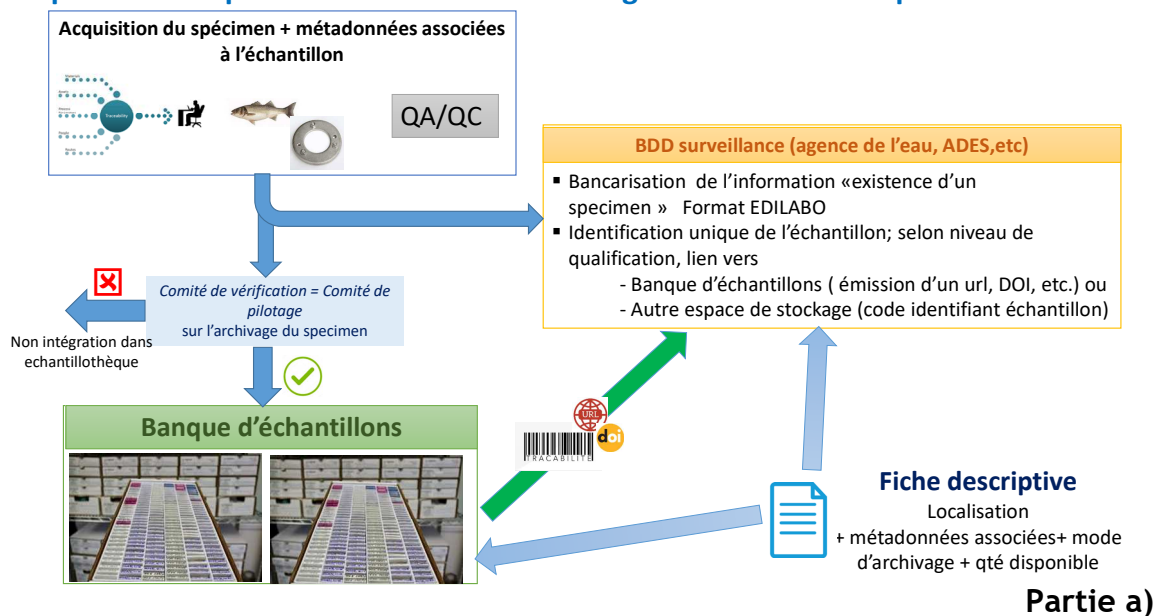


Figure 3 complète la complète pour l'infrastructure spectrothèque.

Acquisition de spécimens en vue de leur intégration dans la banque d'échantillons



Partie a)

Exploitation des données de la banque d'échantillons physiques

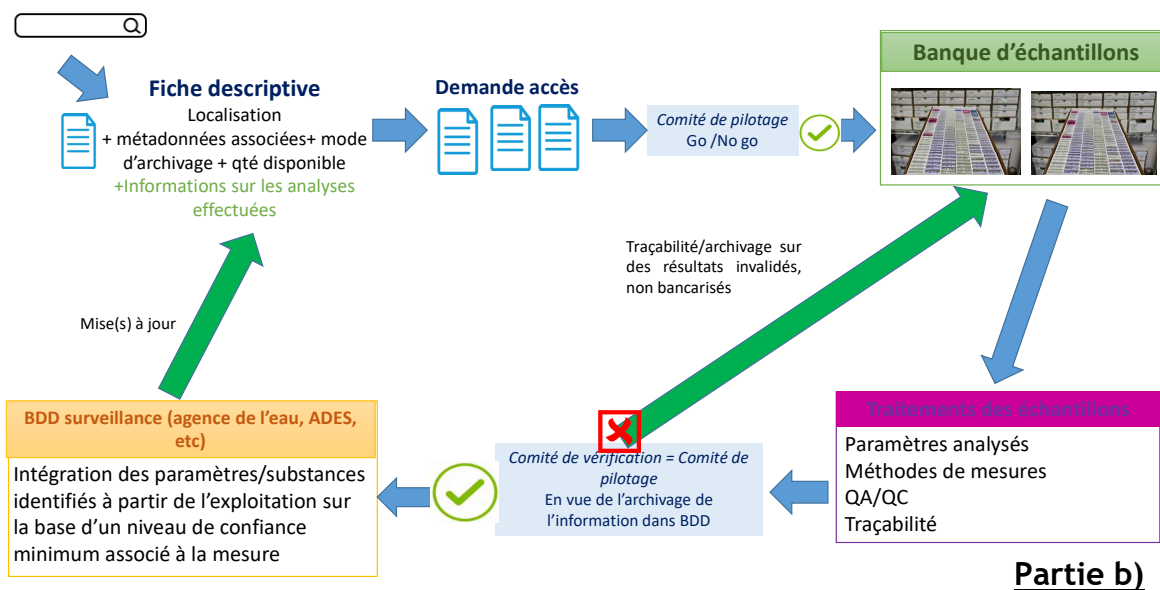
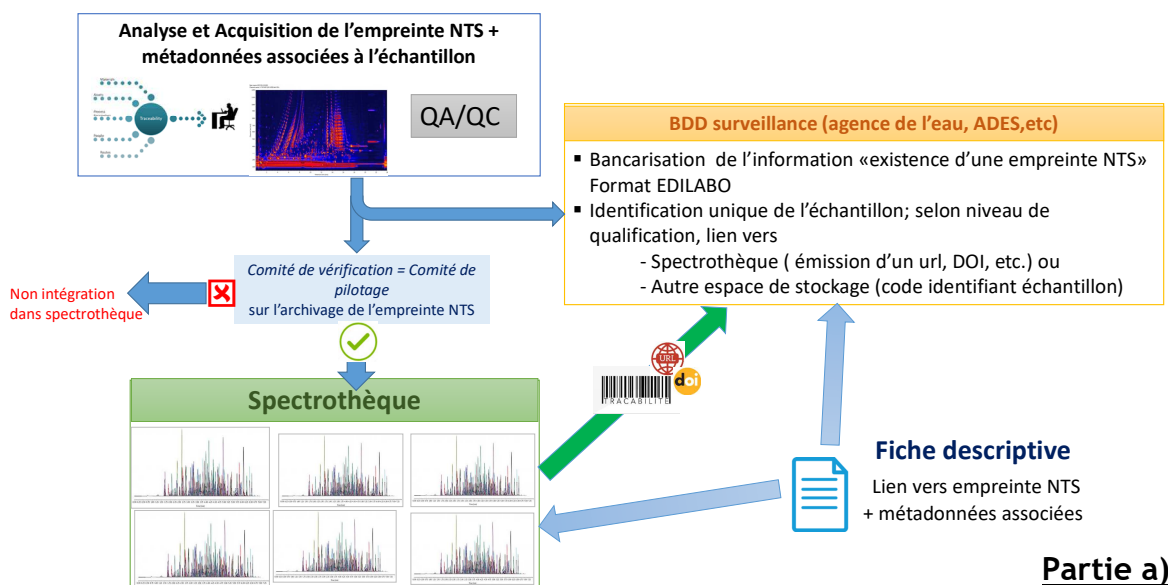


Figure 2 : Banque d'échantillons physiques : illustration du flux de l'information/donnée pour l'acquisition initiale des échantillons physiques de la banque (partie a) et l'exploitation dans le temps de la banque d'échantillons (partie b)

Acquisition initiale en vue d'intégrer la spectrothèque





Consortium scientifique et technique
Région, Paysan, INRS, Info et Us

QUEL MONITORING EN 2050 ET QUELS OUTILS UTILISERONS-NOUS EN 2050?

- ☐ L'évaluation des effets biologiques s'améliore rapidement
- ☐ Nouvelle idée d'ESB virtuels complémentaires basés sur des spectres non ciblés
- ☐ L'évaluation intégrée de l'exposition des données de surveillance humaine et environnementale figure en bonne place dans l'agenda politique
- ☐ Les ESB testent de nouvelles matrices environnementales, par exemple boues d'épuration, particules en suspension et échantillons passifs pour l'eau et l'air
- ☐ Nouveaux compartiments de surveillance, par exemple environnements urbains / mégapoles
- ☐ Le big data /data mining permettra une analyse plus intégrative des paramètres chimiques et biologiques.



2



Consortium scientifique et technique
Région, Paysan, INRS, Info et Us

- Actions à suivre :
- Echange avec NORMAN pour supprimer les difficultés techniques
- Echanges avec utilisateurs (potentiels)?
 - UBA J Koschorreck,
 - W Brack (projet SOLUTIONS)
 - ???
- ?



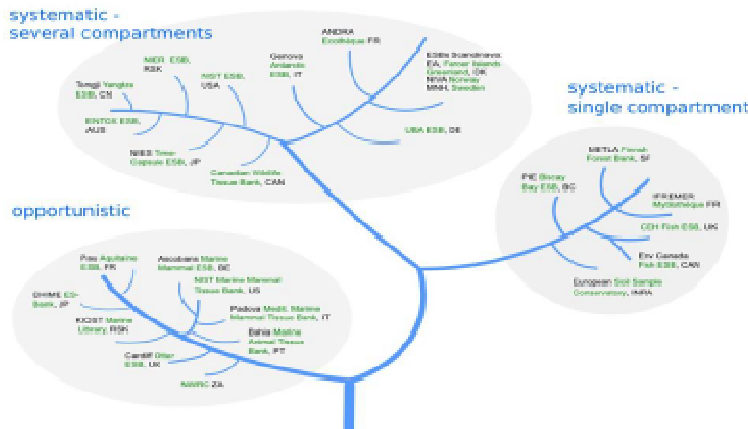
76



Consortium
académique
et technique

FROM IFREMER

ENVIRONMENTAL SPECIMEN BANK – A STATEMENT FOR DIVERSITY





Consortium
académique
et technique

En France

- Mytilothèque: IFREMER
- Ecothèque: ANDRA
- Conservatoire des sols Pédothèque : INRA
- Génols qui s'est développée à partir des sols du RMQS en constituant une ADNthèque: INRA
- Ressources biologiques marines EBMRC France
- Observatoire de Recherche sur la Qualité de l'Environnement du Grand Sud-Ouest Européen
- Service d'Observation et d'Expérimentation pour la recherche sur le long terme en Environnement (SOERE), [ECOSCOPE](#) labellisé en 2012 par [AliEnvi](#).
- Collections patrimoniales, Banque de gènes, Banques de graines : conservatoires botaniques nationaux (CBN)
- Les herbiers (pour le monde végétal): MNHN + l'Université Claude Bernard de Lyon
- ANSES ILNR
- Pau ?

Un foisonnement d'entités/organisations
Statuts/réalité obscures



Consortium
académique
et technique

En France

Accessibilité

- ❖ Quels niveaux d'accessibilité doivent être présents ?
 - ❖ Pour qui (public/privé
 - ❖ Pour quels usages (que R et D /APP ou projets « commerciaux »
- ❖ Comment structurer les accès ?
- ❖ Comment ces projets (de banque ?) sont construits ?



Exigence de transparence et d'information des citoyens



- L'objectif d'ECOSCOPE est de mettre en place un pôle de données d'observation pour la recherche sur la biodiversité afin d'en documenter et comprendre l'état et la dynamique, les mécanismes sous-jacents et construire les scénarios du futur de la biodiversité.
- ECOSCOPE est une infrastructure qui doit contribuer à mieux organiser la collecte, la gestion et la valorisation des données à travers l'accès distant à différents types d'observations de la biodiversité.
- Ces données proviennent d'observatoires de recherche sur la biodiversité, de centres de ressources biologiques et de collections, mais également des données d'expérimentation et restent gérées par les laboratoires.
- Elles sont utilisées pour alimenter la recherche et fournir informations et synthèses pour l'appui à la décision politique.
- Le pôle de données est organisé à partir de l'existant et contribue à l'animation d'une communauté de pratiques entre producteurs de données d'observation de la biodiversité d'une part et producteurs et utilisateurs de données d'autre part.
- ECOSCOPE est un projet multi-organismes. Il est inscrit sur la feuille de route des Infrastructures de Recherche nationales (IR) et a été labellisé Système d'Observation et d'Expérimentation pour la recherche en Environnement (SOERE) par l'Alliance nationale de la recherche pour l'Environnement (AllEnvi) sur la période 2011-2015. Denis Couvet (Muséum National d'Histoire Naturelle) en assure la coordination scientifique et la Fondation pour la Recherche sur la Biodiversité (FRB) la coordination opérationnelle. Les travaux menés à travers ECOSCOPE sont réalisés en lien avec le ministère de la Recherche et les acteurs de la recherche, et en étroite collaboration avec les ministères de l'Environnement et de l'Agriculture ainsi que les acteurs de la société.
 - <http://www.fondationbiodiversite.fr/images/documents/ECOSCOPE/ECOSCOPE.pdf>

Les objectifs du conservatoire

- ☐ Analyser les phénomènes de dérive analytique
- ☐ Construire la mémoire de nos sols
- ☐ Mettre à disposition une banque d'échantillons

- Le Conservatoire est organisé en ateliers.
- Toutes les activités suivent une démarche qualité suivant le référentiel INRA (Mission Qualité, 2005). Selon les programmes concernés, les échantillons subissent différents traitements. Les tâches sont décrites dans des fiches spécifiques (*figure 1*).
- Chaque projet est d'abord décrit dans une «fiche projet», les modalités de préparation des échantillons sont, quant à elles définies dans une «fiche travaux» qui correspond à un lot d'échantillons donné, à une date donnée et pour un programme défini.
- Ainsi, les «menus analytiques» et demandes de préparations peuvent évoluer au cours d'un même projet.
- Un système de codes-barres est utilisé qui permet de lier des échantillons à la base de données, d'actualiser leur appartenance à un nouveau programme, de suivre leur poids actuel et d'en gérer les stocks, et de connaître les analyses qu'ils ont subies ainsi que leur localisation dans la pédothèque.
- Toutes les informations disponibles sur les échantillons sont saisies en base de données DoneSol2
- bien que les coordonnées exactes des prélèvements soient connues, elles sont soumises à la confidentialité.

Mise à disposition des échantillons –exemples d'utilisations

- Des sous-échantillons ou aliquotes sont mis à disposition à la demande, le plus souvent pour des projets de Recherche dans lesquels l'Unité InfoSol est impliquée scientifiquement.
- L'équipe participe, en amont, à l'élaboration du protocole et aux réflexions sur les quantités et la nature des échantillons fournis ou sur les stratégies d'échantillonnage le cas échéant.
- Dans certains cas, leur sélection peut être faite sur des critères physicochimiques, pédologiques, d'occupation des sols, de gradient de teneurs, de localisation géographique, etc. Dans d'autres cas, **c'est le critère d'exhaustivité qui est appliqué**, et l'ensemble des échantillons du RMQS est mis à disposition.
- En aval, l'unité Info-Sol gère les données se rapportant à l'échantillon (anciennes et nouvelles analyses).
- Enfin, l'unité InfoSol participe le plus souvent au traitement, à l'interprétation et à la valorisation des résultats. La partie suivante présente une synthèse des utilisations des échantillons du Conservatoire à ce jour.

- *La plateforme GenoSol de l'UMR Agroécologie de l'INRA de Dijon est une structure unique en Europe, centrée sur la (micro)biologie des sols. Ses missions s'articulent autour de trois activités principales : un Conservatoire des Ressources Génétiques Microbiennes, une Plateforme Technique de caractérisation moléculaire du métagénome microbien, un Système d'Information sur la diversité microbienne des sols et de l'environnement. Depuis 2015, la Plateforme GenoSol est certifiée ISO 9001 pour le Conservatoire des Ressources Génétiques Microbiennes.*
- Objectif : acquérir, stocker et caractériser des ressources génétiques microbiennes des sols.
- Son but ultime est de produire un système de référence basé sur la caractérisation moléculaire des communautés microbiennes du sol qui permettra aux scientifiques d'interpréter des analyses à petite échelle de temps et d'espace.
- Cette collection d'échantillons est disponible pour la communauté scientifique nationale et internationale en termes de ressources génétiques des sols

Accès en construction



AQUA-REF
réseau national
pour la recherche
en aquariologie

La Mytilothèque

Evaluation de l'intérêt de la demande

Examiner la pertinence de cette demande, éventuellement de faire préciser certains points au demandeur avant de donner une réponse définitive. **Le comité de département BE (COBE) est réputé compétent pour cet examen qui sera mis à l'ordre du jour de la première réunion suivant la réception de la demande.**

Les demandes émanant d'ifremer, y compris de BE, devront être examinées avec la même rigueur que les demandes externes

Examen des quantités requises

Les demandeurs devront préciser la **quantité minimale de chaque échantillon** dont ils ont besoin pour réaliser leur étude. La justification de cette quantité sera appréciée par BE et éventuellement négociée.

Nécessité d'établir le ratio Quantité / intérêt

Citations et références

Le demandeur devra donc associer le département BE à la publication des résultats

Préparation d'une partie aliquote

Les échantillons d'origine, constituant le stock de la banque, ne doivent pas sortir du département.

Aliquotage à la demande par des personnels compétents BE

Fiche de sortie

Fiche à compléter

Retour éventuel des reliquats, mise à jour de l'inventaire

Au cas par cas, les reliquats issus de demandeurs connus de BE peuvent être récupérés, surtout s'il s'agit d'échantillons dont il reste très peu en banque.

Après envoi des parties aliquotes, ou retour éventuel des reliquats, le responsable de la banque jugera nécessaire ou non de mettre à jour le fichier d'inventaire, principalement sur le plan des quantités restantes.



AQUA-REF
réseau national
pour la recherche
en aquariologie

Collections MNHN



18

- Bases de données scientifiques du MNHN
 - <https://science.mnhn.fr/institution/mnhn/search>

CHIMIOTHÈQUE ET EXTRACTOTHÈQUE

- La chimiothèque est composée de 1 300 molécules, essentiellement des produits naturels historiques qui lui confèrent une grande diversité et une réelle originalité. Ces produits sont issus des collections du Muséum et des travaux de recherche en cours. Chaque molécule est conservée sous forme de poudre amorphe à température ambiante et en solution dans le DMSO à -20°C.
- L'extractothèque se compose de 240 extraits issus de champignons supérieurs et de spécimens d'invertébrés marins de la zoothèque du Muséum. Ces extraits sont conservés en solution dans l'éthanol et dans le DMSO à -20°C.
- Les produits et extraits naturels en solution dans le DMSO sont conditionnés sous forme de plaque 96 puits facilement accessibles pour les criblages à haut débit. Ces collections sont gérées par l'unité mixte de recherche MCAM (UMR 7245 CNRS/MNHN).

- Demande de prêt et consultation de collections

Formuler une demande de prêt via COLHELPER**Comment formuler une demande de prêt ?**

Toute demande de prêt doit être effectuée via la [plateforme internet Col/Helper](#) pour laquelle une création de compte est obligatoire.

NB. Toutes les demandes de prêt pour exposition et/ou diffusion sont à renseigner dans la rubrique "Exposition". Pour les demandes de prêts scientifiques, elles doivent être renseignées dans la rubrique correspondant à la collection concernée.

Dans le cadre d'exposition et de diffusion, toute demande officielle doit être effectuée dans un délai minimum de 4 mois avant l'ouverture de l'exposition. Ce délai a été défini en fonction des ressources disponibles au bureau de la régie, des interactions avec l'emprunteur et des contraintes réglementaires de circulation des biens concernés.

Le bureau de la régie se réserve le droit de ne pas accepter ces demandes si le délai est trop court au vu de la demande.

La gestion des demandes de prêts à but de diffusion est assurée par le bureau de la régie du muséum qui constitue le seul interlocuteur direct pour toutes les questions relatives à la demande d'emprunt.

L'Observatoire pérenne de l'environnement (OPE) a pour objectif de faire un état initial de l'environnement. Il est unique en France car il est le premier observatoire à étudier simultanément l'ensemble des compartiments de l'environnement dans tous les milieux d'un vaste territoire d'étude. Pour mener à bien sa mission, il dispose de huit stations de mesures et d'une écothèque

L'Écothèque est une bibliothèque d'échantillons environnementaux. Des prélèvements de toute nature seront préparés, conditionnés et **conservés pendant au moins 100 ans**.
 L'objectif est de préserver la mémoire de la qualité de l'environnement d'aujourd'hui pour être capable de l'analyser demain, avec des méthodes qui seront peut-être plus sensibles, ou pour y rechercher des éléments qui n'ont pas été analysés aujourd'hui.

- EN COURS DE CONSTRUCTION


17



**Centre National de ressources biologiques marines-
EMBRC**

- EMBRC-France met en accès une variété de modèles biologiques marins ainsi qu'un environnement technologique de haut niveau pour leur exploration fonctionnelle. L'infrastructure développe les modèles visant à compléter l'arbre du vivant ainsi que les ressources génétiques et génomiques afférentes.
- EMBRC-France est construit autour des services suivants :
 - ✓ Accès aux modèles biologiques marins, couvrant toutes les lignées de l'arbre du vivant ;
 - ✓ Accès à des dispositifs d'expérimentation ex situ, pour l'élevage ou la culture ;
 - ✓ Accès aux moyens logistiques pour le génotypage et le phénotypage de ces modèles (-omiques, imagerie, performances en culture ...) ;
 - ✓ Accès à des ressources génétiques de certains organismes modèles, procaryotes ou eucaryotes ;
 - ✓ Accès à des ressources numériques sur les organismes et les écosystèmes marins

<http://www.embrc.eu/>


18

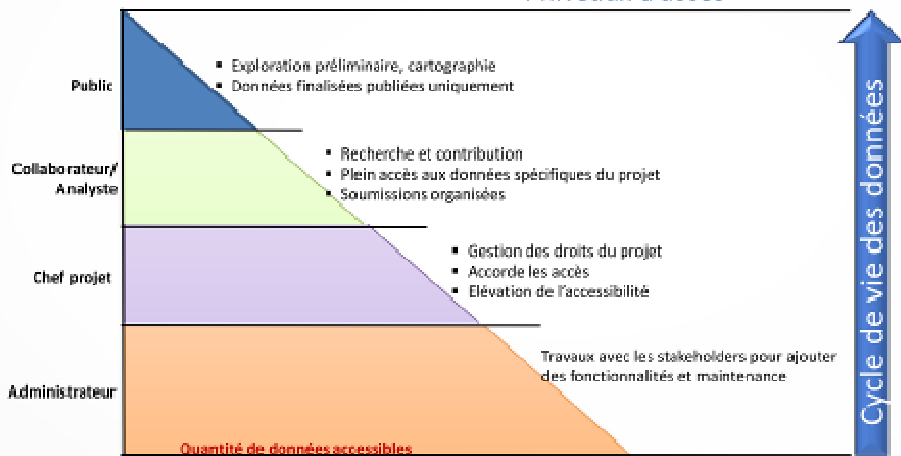


Conseil scientifique
et technique

BRGM, IFREMER,
INRS, IRD et UPR

Ex NIST STAR

4 niveaux d'accès



Cycle de vie des données



Conseil scientifique
et technique

BRGM, IFREMER,
INRS, IRD et UPR

NIST-NMMTB

☐ En raison des coûts associés à la collecte et à la conservation à long terme des échantillons et à la valeur croissante des échantillons après des années (voire des décennies) de stockage, une politique doit être élaborée pour déterminer l'accès à ces échantillons. Cependant, avant de définir une politique d'accès appropriée, il convient de répondre aux questions suivantes: Quel est le but de la banque d'échantillons?

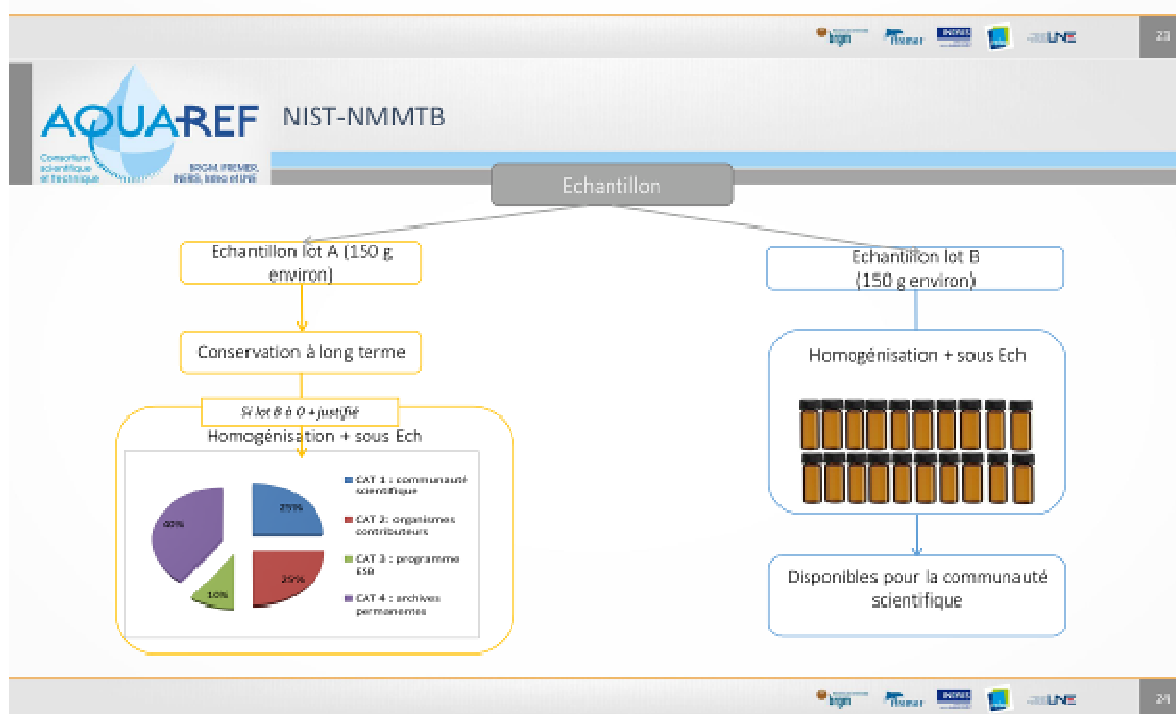
☐ La banque a-t-elle pour objectif de fournir une banque de spécimens à la disposition de tous les programmes désirant des échantillons à des fins de suivi ou une collection "à long terme" d'échantillons pour des analyses rétrospectives qui ne peuvent être réalisées qu'avec des échantillons stockés?

☐ Généralement, cette dernière affirmation est vraie pour la plupart des programmes de banque d'échantillons.

☐ Les demandes d'échantillons de la banque doivent être considérées en fonction de plusieurs critères.

- ☐ Premièrement, les demandes doivent justifier la nécessité d'utiliser des échantillons de la banque, c'est-à-dire qu'un spécimen en banque est-il nécessaire pour atteindre l'objectif visé ou est-ce que les spécimens recueillis peuvent maintenant servir le projet?
- ☐ Deuxièmement, les procédures de contrôle qualité doivent être mises en place par ceux qui demandent les échantillons pour assurer la qualité des résultats d'analyse obtenus avec les échantillons stockés.

Enfin, le demandeur doit accepter de fournir tous les résultats d'analyse au programme de la banque de spécimens pour inclusion dans la documentation de l'échantillon.



Le NMMTB n'est pas destiné à être utilisé comme source facilement accessible de tissu de mammifère marin pour tout chercheur ayant besoin d'échantillons, mais uniquement pour la recherche nécessitant des échantillons stockés du passé. C'est un point important, en particulier pour les tissus de mammifères marins, en raison des restrictions liées à l'acquisition de tels tissus, car certains de ces animaux sont considérés comme des espèces en danger de disparition.

- Catalogue en ligne des échantillons bancarisés
- Demande en ligne d'accès
- Motivation de la demande et notamment justification précise du besoin d'accès à des archives et pas uniquement à des échantillons
- Motivation des quantités demandées pour échantillons physiques
- Politique d'accessibilité
- Accessibilité graduelle selon profil utilisateurs
- Évaluation par comité experts
- Participation aux coûts d'envoi
- Paiement à l'échantillon
- Demande toute l'année ou appel annuel
- Exigence de citation dans la valorisation
- Exigence que les résultats soient retournés vers la base pour meilleure connaissance des spécimens
- Exigence de retour des échantillons non utilisés

- Questions du cout d'accès :
 - | Soit intégration en tant que partenaire : cout « minime » pour l'entretien et la construction de la base
 - | Soit sans partenariat: cout plus élevé (approche plus commerciale..)
- La notion d'intégration dans le partenariat permet de mieux verrouiller l'utilisation, l'interprétation des données
- Notion très importante de retour de l'information produite!! Ce sera la richesse d'une banque pas seulement faciliter l'accès à des échantillons (cela me semble assez commercial...) mais plutôt augmenter la connaissance et permettre une acquisition croisée de connaissance (notion de data mining précédemment évoquée...)
 - | Besoin de discuter de la PI des données acquises...

- Gestion des demandes
- toutes les plateformes fonctionnent:
 - | À partir d'une entité existante (INRA, MNRH, IFREMER...)
 - | Ont un comité d'évaluation de la demande qui s'appuie sur des critères scientifiques (pertinence, nécessité...) et pratiques (quantité nécessaires)
 - | ont un droit de regard sur les résultats (sauf qd que commercial?)
- | Notion de mise en réseau des banques
 - Besoin d'une Interface qui puisse faire des liens entre les bases:
 - liaison géographique GIS ?
 - ...???



CONSEIL NATIONAL
DE L'EAU
ET DE L'ENVIRONNEMENT

Et après

- Organiser une réunion en T1 2019 pour échanger par les plus grosses entités en Fr :
 - | EMBRC-France Nicolas PADE
 - | INRA Genosol+ ADNthèque
 - | MNHN
 - | IFREMER
 - | AFB
 - | U Pau ?
 - | ANDRA?

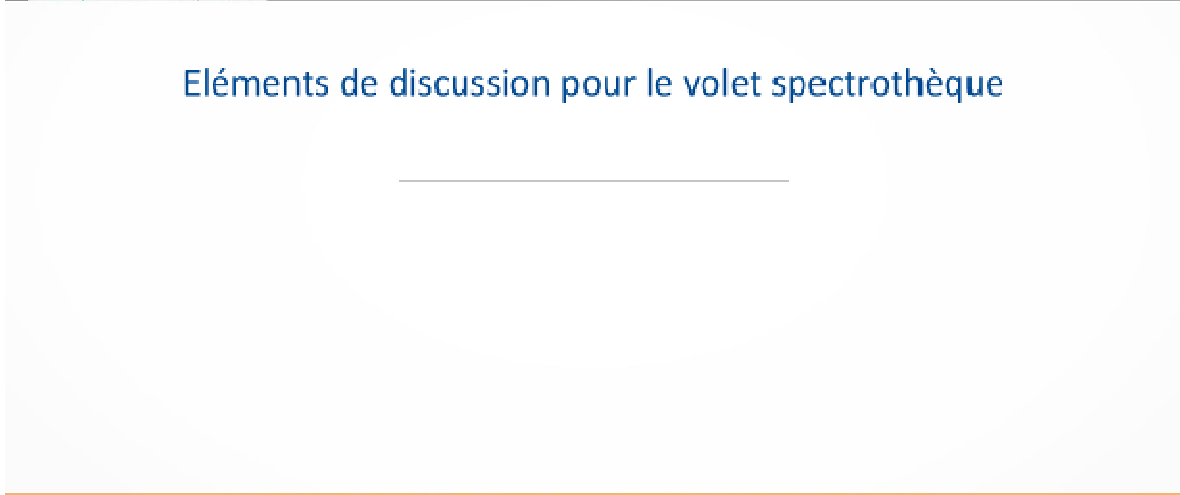


CONSEIL NATIONAL
DE L'EAU
ET DE L'ENVIRONNEMENT



29

Eléments de discussion pour le volet spectrothèque



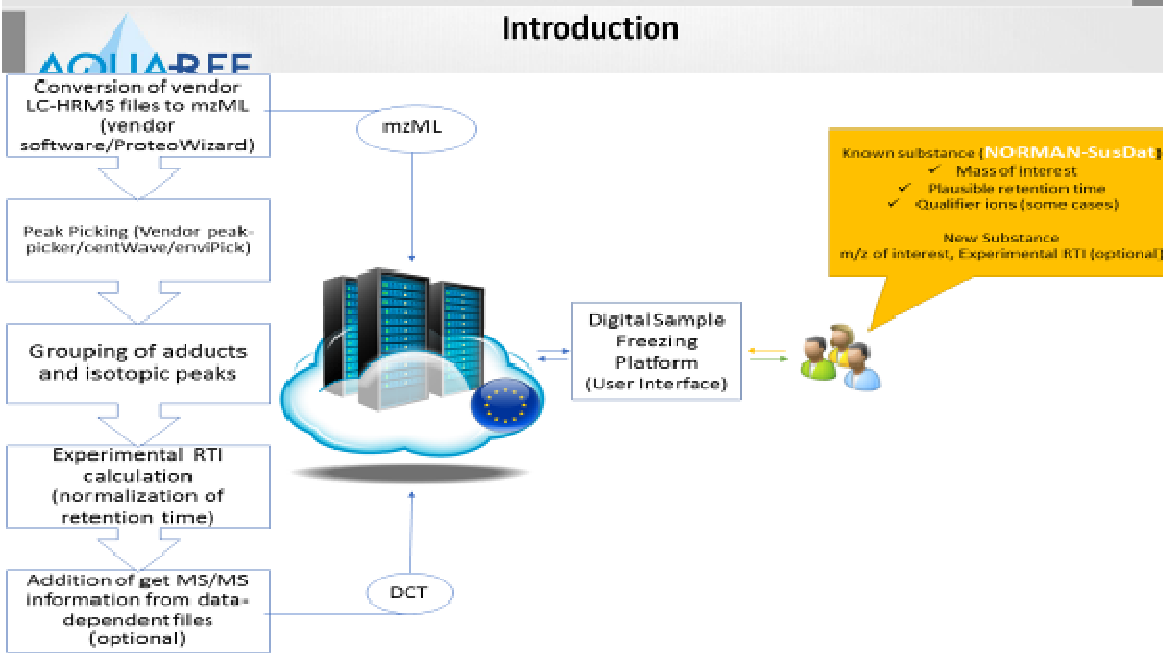


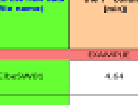
30

Contents

- **Présentation technique DSFP (informations Nikiforos Alygizakis, Jaroslav Slobodnik**
- **Avantages/ limitations techniques et fonctionnelles**
- **Présentation banque BRGM**
- **Avantages / limitations techniques et fonctionnelles**

Introduction





Consortium
scientifique
de l'écotoxicité

Description of NTS DCT

Sample mass spectral information (MS1 information)



NTS DCT

Example identification (check this box first the name)	Retention time in the 1 st column [min]	Retention time in the 1 st column [min]	Mass of ion [m/z] (peak or component)	Intensity of the ion	Intensity of the ion in the blank	Ion type		MSEMS available	Category	Proposed identification (name of the substance as it is for NTS identification)	Molecular formula	Exact Mass
						Other						
EXAMPLE	EXAMPLE	EXAMPLE	EXAMPLE	EXAMPLE	EXAMPLE	EXAMPLE	EXAMPLE	EXAMPLE	EXAMPLE	EXAMPLE	EXAMPLE	EXAMPLE
Observed	4.54		255.1012	1666665	0	NA		Yes	Target	255.1012	C ₁₂ H ₁₅ N ₃ O ₄	
Decomposed	16.68		243	14222	0	Other	Base	No	Unknown	n/a	n/a	
YOUR DATA						YOUR DATA						

Echanger avec Valeria
: je ne comprends pas exactement ce qui sera géré en
automatique et ce qu'il faudra rentrer composés par composés !
(exemple level of confirmation??

MOVING SAMPLE	CAT. No.	CA	Reference standard (ratio of MS/MS)	Reference standard (ratio of MS/MS)	Reference standard (ratio of MS/MS)	Reference standard (ratio of MS/MS)	Reference standard (ratio of MS/MS)	Reference standard (ratio of MS/MS)	Reference standard (ratio of MS/MS)	Reference standard (ratio of MS/MS)	Reference standard (ratio of MS/MS)	Reference standard (ratio of MS/MS)
CONFINED PROPOSE	1012-249	0.2	characteristic pattern at ion 243 shows presence of chlorine atom	345-332-15.3, 356-345-15.5, 380-367-16.0	3804.18	11/12/2018	12/13/2018	12/13/2018	12/13/2018	12/13/2018	12/13/2018	12/13/2018
n/a	n/a	0.18										
YOUR DATA						YOUR DATA						

LC-MS/MS - Data attached

New chromatogram: Positive-negative mode MS-MS, (Organization: **abundance**, Ionization mode: **FOUR**, Collision Energy in eV or %: **Instrument**, Matrix: **Scanning Time, Country, Date of Sampling** (YYYY-MM-DD), Project abbreviation, Origin: **Sample Data**)

MS/MS	Intensity	Retention time	Reference standard (ratio of MS/MS)	Reference standard (ratio of MS/MS)
1012-249	0.2	1012-249	characteristic pattern at ion 243 shows presence of chlorine atom	345-332-15.3, 356-345-15.5, 380-367-16.0

GC-MS - Data attached

New chromatogram: Positive-negative mode MS-MS, (Organization: **abundance**, Ionization mode: **FOUR**, Collision Energy in eV or %: **Instrument**, Matrix: **Scanning Time, Country, Date of Sampling** (YYYY-MM-DD), Project abbreviation, Origin: **Sample Data**)

MS/MS	Intensity	Retention time	Reference standard (ratio of MS/MS)	Reference standard (ratio of MS/MS)
1012-249	0.2	1012-249	characteristic pattern at ion 243 shows presence of chlorine atom	345-332-15.3, 356-345-15.5, 380-367-16.0



Consortium
scientifique
de l'écotoxicité

Description of NTS DCT

Sample mass spectral information



NTS DCT

Mass of ion [m/z]	Retention time [min]	Intensity of the ion	File
66.11185	0.319	6116	D:\POS_25_LC-ESI-QTOF_Effluent wastewater Stuttgart Muhlhausen_ Stuttgart_Germany_03.05.2018_1TEN ANSWER_28543.mvML
74.09039	0.319	15452	D:\POS_25_LC-ESI-QTOF_Effluent wastewater Stuttgart Muhlhausen_ Stuttgart_Germany_03.05.2018_1TEN ANSWER_28543.mvML
320.97146	0.319	5356	D:\POS_25_LC-ESI-QTOF_Effluent wastewater Stuttgart Muhlhausen_ Stuttgart_Germany_03.05.2018_1TEN ANSWER_28543.mvML
216.99228	0.802	2948	D:\POS_25_LC-ESI-QTOF_Effluent wastewater Stuttgart Muhlhausen_ Stuttgart_Germany_03.05.2018_1TEN ANSWER_28543.mvML
357.02192	0.899	1499	D:\POS_25_LC-ESI-QTOF_Effluent wastewater Stuttgart Muhlhausen_ Stuttgart_Germany_03.05.2018_1TEN ANSWER_28543.mvML
359.12252	0.898	1736	D:\POS_25_LC-ESI-QTOF_Effluent wastewater Stuttgart Muhlhausen_ Stuttgart_Germany_03.05.2018_1TEN ANSWER_28543.mvML
128.06199	0.909	1300	D:\POS_25_LC-ESI-QTOF_Effluent wastewater Stuttgart Muhlhausen_ Stuttgart_Germany_03.05.2018_1TEN ANSWER_28543.mvML
129.01806	0.936	3512	D:\POS_25_LC-ESI-QTOF_Effluent wastewater Stuttgart Muhlhausen_ Stuttgart_Germany_03.05.2018_1TEN ANSWER_28543.mvML

Pour chaque échantillon = l'information sur ce qui
a été détecté

Composés traceurs (SemiQuantification)

Spiked Compound Name	SAMPLES	Concentration level (units)	Response expression (Peak Area or Maximum Intensity)	Concentration level 1	Response at concentration level 1	Concentration level 2	Response at concentration level 2	Concentration level 3	Response at concentration level 3	Concentration level 4	Response at concentration level 4
Example	Example	Example	Example	Example	Example	Example	Example	Example	Example	Example	Example
				Example	Example	Example	Example	Example	Example	Example	Example
Clarithromycin	(C ₁₂ H ₁₇ ClO ₆)	mg/L	Peak Area	5.5	123.55	10	123.55	15	123.55	20	123.55
Ciprofloxacin	(C ₁₇ H ₁₄ F ₂ N ₂ O ₃)	ng/L	Peak Area	10	123.55	20	123.55	30	123.55	40	123.55

NORMAN Digital Sample Freezing Platform					Mass Page	Batch mode	Contribute	More +
Basic Information	Sample Metadata	Instrumental Metadata	RTI Calibration	Upload Files				
Who is contributing					Informations sur le laboratoire			
All fields are compulsory								
Institute Name								
EI - Environmental Institute								
Organization City								
Kos								
Organization Country								
Slovakia								
Laboratory performing the analysis								
Environmental Institute								
City of the laboratory performing the analysis								
Kos								
Country of the laboratory performing the analysis								
Slovakia								
Contact person family name								
Slobodnik								
Contact person first name								
Jozsef								
What is contributing					Informations sur l'analyse			
All fields are compulsory								
Instrument								
LC-ESI-QTOF								
Short name describing the sample								
Danube J0883								
Title of Project (Acronym)								
Joint Danube Survey 3								
Country from which the sample was taken								
Romania								
City from which the sample was taken								
Tulcea								
Sampling Date in dd-mm-yyyy								
08-03-2016								
Analysis Date in dd-mm-yyyy								
28-04-2016								
How many times has the sample been preconcentrated (Preconcentration factor)								
40000								
Unique Identifier								
R042B16								
Proceed to step 3/5								

NORMAN Digital Sample Freezing Platform
Main Page
Batch mode
Contribute
More

Basic information
Sample-Metadata
Instrumental-Metadata
RTI_Combination
Upload-Files

Informations sur l'échantillon

General Sample Information

Longitude in Decimal (Compulsory)

20.5242

Latitude in Decimal (Compulsory)

45.1935434

Precision of coordinates (Compulsory)

Average (range 10-100 m)

Type of data source

Surveys

Type of monitoring

Investigative

If sample was taken from border among 2 countries, which is the 2nd country

Specify any remarks or important extra information regarding the samples

Type of matrix (Compulsory)

Choose matrix from the list



Consortium scientifique de l'écologie
Normandie, Île de France, Bretagne, Pays de la Loire, Occitanie, Provence-Alpes-Côte d'Azur, Rhône-Alpes, Auvergne-Rhône-Alpes, Nouvelle-Aquitaine, Grand Est, Bourgogne-Franche-Comté, Centre-Val de Loire, Hauts-de-France, Île-de-France, Normandie, Occitanie, Provence-Alpes-Côte d'Azur, Rhône-Alpes, Auvergne-Rhône-Alpes, Nouvelle-Aquitaine, Grand Est, Bourgogne-Franche-Comté, Centre-Val de Loire, Hauts-de-France, Île-de-France

Type of matrix (Compulsory)

Municipal

Fraction (Compulsory)

Collected fraction

Type of waste water (Compulsory)

Waste water - Municipal

Proag. processes

NAME of river basin

Specify

Additional Information

pH

7.4

Temperature (°C)

10.4

Dissolved organic carbon (mg/l)

30

Capacity by population equivalent (Compulsory)

100000

Daily flow potential (Compulsory)

100000

Type of treatment plant (Compulsory)

0 - Evaporation + primary sedimentation tank + activated sludge (including SBR) 1 - 4 - 5 - 6

Type of tertiary treatment (Compulsory)

0 - No tertiary treatment + noxious substances - chlorination

MS & Dry matter content (mg/l)

MS (mg/l)

MS (mg/l)

Volume of the reactor (m³)

Proceed to step 05

Informations sur la STEP ?

Valeria c'est obligatoire???



NORMAN Digital Sample Freezing Platform Main Page Batch mode Contribute More ▾

Basic-Information Sample-MetaData Instrumental-MetaData RTI_C Informations sur l'analyse

☐ I have positive ionization for the same samples

☐ I have negative ionization for the same samples

Proceed to step 4/5

Basic-Information Sample-MetaData Instrumental-MetaData RTI_Calibration Upload-Files

☒ I have positive ionization for the same samples

All fields are compulsory

Instrument manufacturer
Bruker

Model type and number
emulo Impact

Analytical column
Thermo Accell HPLC C18

Column dimensions (length mm; ID, mm; Particle size um)
12.2um, 2.1x100mm

Injection volume
10

Column temperature (°C)
25

Composition of the mobile phase
A: RT 15 min (method with 0.1% BSA) B: RT 15 min (method with 0.1% BSA) C: RT 15 min (method with 0.1% BSA) D: RT 15 min (method with 0.1% BSA)

NORMAN Digital Sample Freezing Platform Main Page Batch mode Contribute More ▾

Basic-Information Sample-MetaData Instrumental-MetaData RTI_Calibration Upload-Files

☐ I have not injected these these calibrants OR I am using GC-HRMS

You have to input the RT of the calibrants. Let the cells empty for compounds without RT.

	Name	RT	CAS	Formula	Ion
1	Guanylfurea	1.325	141-83-3	C2H6N4O	103.0514
2	Amitrol	1.392	61-82-5	C2H4N4	85.0509
3	Histamine	1.642	51-45-6	C5H9N3	112.0859
4	Chlomequat	1.875	999-81-5	C5H13ClN	122.0731
5	Methamidophos	2.625	10265-92-5	C2H8NO2P3	142.0086
6	Vancomycin	3.192	1404-90-6	C66H75Cl2N9O24	724.7224
7	Cefoperazone	4.342	62893-19-0	C25H27N9O8S2	646.1497
8	Trichlorfon (Dylox)	5.142	52-68-6	C4H8Cl3O4P	256.9299
9	Butocarboxim	5.992	34681-10-2	C7H14N2O2S	191.0849
10	Dichlorvos	6.908	62-73-7	C4H7Cl2O4P	220.9532
11	Tylosin	7.975	1401-69-0	C46H77NO17	916.5264
12	TCMTB	9.208	21564-17-0	C9H6N2S3	238.9766
13	Rifaximin	10.025	80621-81-4	C43H51N3O11	786.3596
14	Spinosad A (Spinosyn A)	11.525	131929-60-7	C41H65NO10	732.4681
15	Emamectin B1a	12.475	121124-29-6	C49H75NO13	886.5311
16	Avermectin B1a (Abamectin)	13.674	71751-41-2	C48H72O14	890.5260
17	Nilgericin	13.908	28380-24-7	C40H68O11	725.4834
18	Ivermectin B1a	14.458	70288-86-7	C48H74O14	892.5436

Proceed to step 5/5

Utilisation des RTI index pour aide sur les inconnus

AQUAREF Compound selection

Substance name or CAS or SMILESKey
compound

OR

Procurer file
204.000000

Substance name or CAS or SMILESKey
[ALADPTXXXPTUT-UHFFFAQYSA-N]

Ionization (☐ Unlock)
Positive

Compound can be ionized positively and negatively!
Adduct (☐ Unlock)
[M+H]⁺

Possible adducts: [M+H]⁺, [M+Na]⁺
Predicted RT: Positive
[M+H]⁺ 150.000000

Reset

Substance name or CAS or SMILESKey
Trimethoprim [73-57-0] [NEDVHCEMGRBQM-UHFFFAQYSA-N]

Ionization (☐ Unlock)
Positive

Adduct (☐ Unlock)
[M+H]⁺

Spectral information available
Predicted RT: Positive
[M+H]⁺ 258.000000

Reset

Substance name or CAS or SMILESKey
Perfluorooctane sulfonic acid (PFOS) [1763-23-1] [YFBSUTJLHUPNCE-UHFFFAQYSA-N]

Ionization (☐ Unlock)
Negative

Adduct (☐ Unlock)
[M-H]⁻

Spectral information available
Predicted RT: Negative
[M-H]⁻ 410.000000

Reset

Compounds with prior-knowledge on spectral behavior:

- ✓ Ionization pre-selection
- ✓ Adduct pre-selection
- ✓ Qualifier fragments in evaluation of chromatograms module

Log In Home About Us Contact Us

Contributed Samples Results Chromatograms Interactive Map Help

new 10 entries

Select per page

country matrix Project

	country	matrix	Project
152	USA, LC-ESI-QTOF, Seawater GE01_Back Sea_Georgia_25.05.2016_EMBLAS IL_17801.xlsx	Georgia	Water-Surface water-Territorial (marine) water
151	USA, LC-ESI-QTOF, Seawater GE02_Back Sea_Georgia_25.05.2016_EMBLAS IL_17802.xlsx	Georgia	Water-Surface water-Territorial (marine) water
150	USA, LC-ESI-QTOF, Seawater GE03_Back Sea_Georgia_31.05.2016_EMBLAS IL_17803.xlsx	Georgia	Water-Surface water-Territorial (marine) water
149	USA, LC-ESI-QTOF, Seawater GE04_Back Sea_Georgia_31.05.2016_EMBLAS IL_17804.xlsx	Georgia	Water-Surface water-Territorial (marine) water
148	USA, LC-ESI-QTOF, Seawater GE05_Back Sea_Georgia_30.05.2016_EMBLAS IL_17805.xlsx	Georgia	Water-Surface water-Territorial (marine) water
147	USA, LC-ESI-QTOF, Seawater GE07_Back Sea_Georgia_30.05.2016_EMBLAS IL_17806.xlsx	Georgia	Water-Surface water-Territorial (marine) water

Column visibility Copy CSV Excel Print Show 150 entries

Visualisation résultat pour un composé

Select per page

Retention Time [min]	Mass of ion [m/z]	Intensity	MS/MS available	Proposed substance	Level of confirmation of Identification	RTI	Qualifier Ions [m/z] (Intensity)	
All				?	All		All	
UsA_LC-ESI-QTOF_Sediment GE01_Black Sea_Georgia_28.05.2016_EMBLAS IL_15223	7.55	505.2094	1820	No	Chlorhexidine	Exact mass Plausible RT and 9 fragments	477	336.1703/7.61/4295 201.1520/7.59/2120 184.1550/7.61/4528 178.0166/7.61/1476 170.0473/7.61/11056 159.1599/7.61/904 153.0210/7.61/8428 138.0260/7.61/2840 135.1073/7.61/7658
UsA_LC-ESI-QTOF_Sediment GE02_Black Sea_Georgia_28.05.2016_EMBLAS IL_15224	7.55	505.2103	2640	No	Chlorhexidine	Exact mass Plausible RT and 12 fragments	477.1	353.1957/7.61/4620 236.1696/7.61/4592 195.0425/7.61/1185 184.1547/7.61/3728 178.0160/7.59/585 170.0469/7.60/11520 167.1275/7.62/1368 159.1593/7.63/852 153.0206/7.60/8338 138.0260/7.61/2868 135.1073/7.60/7658 100.1134/7.63/3820
UsA_LC-ESI-QTOF_Sediment GE05_Black Sea_Georgia_30.05.2016_EMBLAS IL_15226	7.47	505.2108	3544	No	Chlorhexidine	Exact mass Plausible RT and 13 fragments	470.7	455.1832/7.46/1416 353.1956/7.46/6008 336.1465/7.48/5000 201.1818/7.48/1812 195.0427/7.48/1596 184.1550/7.46/5448 170.0157/7.46/512 170.0473/7.46/17060 167.1283/7.48/2612 159.1596/7.48/1520 153.0206/7.48/10072 138.0266/7.48/3480 135.1077/7.48/8036



NORMAN Digital Sample Freezing Platform

Main Page Batch mode **Registration** Mass

Create Results Visualisation

Step 1: Select compound

You selected 0 compound.

Compound lists from Suspect list exchange

Nothing selected

SEARCH/SELECT SUSPECT 3438

Show 10 entries

Name	Formula	CAS number	InChIKey	ChemSpiderID
1 Butylhexanoate	C10H20O2	CAS RN: 103-48-8	QKQJXUJLHBPQZ-UHFFFAOYSA-N	88200
2 Butylhexylacrylate	C18H34O2	CAS RN: 80-32-0	XOMHUPDQWFOO-UHFFFAOYSA-N	4382
3 Butylacrylate	C7H12O2	CAS RN: 87-67-0	BRBQDQJHSDQD-UHFFFAOYSA-N	8133
4 Butylacrylate	C7H12O2	CAS RN: 137-79-7	QKQJXUJLHBPQZ-UHFFFAOYSA-N	8134
5 Butylacrylate	C7H12O2	CAS RN: 144-82-1	VACQJXUJLHBPQZ-UHFFFAOYSA-N	8137
6 Butylacrylate	C7H12O2	CAS RN: 729-89-7	CYFLXJLHBPQZ-UHFFFAOYSA-N	12561
7 Butylacrylate	C7H12O2	CAS RN: 1337-39-9	PDDQJXUJLHBPQZ-UHFFFAOYSA-N	3142
8 Butylacrylate	C7H12O2	CAS RN: 6889-19-0	QKQJXUJLHBPQZ-UHFFFAOYSA-N	545849
9 Butylacrylate	C7H12O2	CAS RN: 35287-89-3	UJQJXUJLHBPQZ-UHFFFAOYSA-N	371186
10 Butylacrylate	C7H12O2	CAS RN: 8428-12-8	VICTQJXUJLHBPQZ-UHFFFAOYSA-N	78767

Showing 1 to 10 of 40,655 entries

PREVIOUS 1 2 3 4 5 ... 4066 NEXT

Step 2: Select parameters

Step 3: Select samples that you want to screen

Step 4: Submit your request

[illegible]

Visualisation des résultats
Sortie excel par échantillon

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
Compound	Formula	CAS	SMILES	InChI	InChIKey	Identical	Predicted	EL_LC-ESI-	EL_LC-ESI-	EL_LC-ESI-	EL_LC-ESI-	EL_LC-ESI-	EL_LC-ESI-	EL_LC-ESI-	EL_LC-ESI-	EL_LC-ESI-
1	Compound	Formula	CAS	SMILES	InChI	InChIKey	Identical	Predicted	EL_LC-ESI-	EL_LC-ESI-	EL_LC-ESI-	EL_LC-ESI-	EL_LC-ESI-	EL_LC-ESI-	EL_LC-ESI-	EL_LC-ESI-
2	Sulfapyridine	C11H11N3	144-83-2	C1(S)(Nc2cnc3c2ccn3)N	GECHUMH	Mass accu	FALSE	30768	5176	1564	7560	59500	19996	78068	19972	20460
3	ciclotacilin	C15H22N4	3485-34-1	CC1(C)SC2InChI=15/4HGRUNBBI	Mass accu	TRUE	971748	182264	200808	766504	15044	40684	117208	1200228	565752	
4	[3R,3R,4R, C31H41N5	56291-56	CCN(C)C@H	InChI=15/4CDBUQAM	Mass accu	FALSE	102504	598112	58384	102552	152176	1364556	181648	140212	1362634	
5	linezolid	C16H20FN	165-800-03	CC(CO)NC	InChI=15/4TYZROVQI	Mass accu	FALSE	28452	23616	2884	15336	144812	5944	34328	32652	6248
6	Sulfameth	C10H11N3	721-66-6	C1(S)(Nc2cnc3c2ccn3)N	JKIGFTWJ	Mass accu	FALSE	17524	1008	2020		1148	2944	3268	2624	
7	Sisomicin	C19H27N5	323-85-11	C1C@H	InChI=15/4BRWAWWJ	Mass accu	TRUE	7408		10040	65524	25912/242	33420/320	61844/276	54480/277	13108
8	Sparganin	C17H27N7	809-02-43	C1CCNC	InChI=15/4GDVNULLJ	Mass accu	TRUE	24604		8784	21200	8208	15100	9008		12032
9	Trimethop	C14H18N4	738-70-5	C1(Cc2cnc3c2ccn3)N	IE DV/HCEI	Mass accu	FALSE	28728	5680		42312	44404	16472	11632	44676	
10	Corydopel	C10H13N5	73-03-0	C1NC	InChI=15/4OFEZSAMJ	Mass accu	TRUE	9200		9984	40664	18756		38000	57084	
11	Azithromy	C15H17N5	839-05-01	CC(C@H)	InChI=15/4MQTOSIV	Mass accu	FALSE	24264	2996		7784	7380		11124	15468	
12	Clarithrom	C18H19NO	811-03-11	CC(C@H)	InChI=15/4AGQYDEPI	Mass accu	FALSE	34116	14160		78272	60784	18660	12372	14696	
13	2-methyl- C4H5N	2682-20-4	Cn3ccc1c	InChI=15/4BEGLCMHJ	Mass accu	TRUE		8704	5428	10136	3088	2684	10044			
14	Chuangxin	C12H11N	633-39-68	C1C@H	InChI=15/4DKHFLDK	Mass accu	TRUE	4468	2756	2000		4144	1936	2388		
15	Mefenitall	C13H12N2	6315-71-5	C1C@C	InChI=15/4CBUAVAK	Mass accu	TRUE	3432			1116	3548	2772	7052	1896	
16	Elthambut	C10H24N2	74-55-5	NC(C@H)	InChI=15/4AEUTYOVY	Mass accu	FALSE	3272			1696	16580	3916	6704	1284	
17	clindamys	C18H33ClN	183-23-44	CC(C)C@H	InChI=15/4KBUVYVY	Mass accu	FALSE	9036	2308			10944		5184	5424	
18	N-Methyl- C7H13NO	713-63-0	CN1C(C)C	InChI=15/4AAKDPDF	Mass accu	TRUE	11896			6132	6992	3076	11428	5296		
19	Sulfameth	C12H14N4	57-68-1	C1(S)(Nc2cnc3c2ccn3)N	ASWYTON	Mass accu	FALSE	12894	3184			3216	1420		3408	
20	Myxophen	C17H20O6	242-89-33	C12(C)S(C)	InChI=15/4HPN5F5B2	Mass accu	FALSE				3390		3992		14084	13132
21	Ketoconaz	C26H38ClN	652-77-42	CC(CO)N1	InChI=15/4XMYAWYJ	Mass accu	FALSE	1816				1128	1704		898	864
22	Emertodol	C44H76O1		C=C(C)C	InChI=15/4BKZOUCVI	Mass accu	TRUE	1968				3156	94524	10484	3808	14956
23	3,4,3-Pipe	C8H13NO	151-39-36	CC(C@H)	InChI=15/4UXBIFEVIS	Mass accu	TRUE				3284	2776		3888		
24	Dacarbazin	C8H10BrN2	691-98-0	CN(C)N=C	InChI=15/4PCKXTQMI	Mass accu	TRUE		274320	31088						67416
25	Fuzilicillin	C25H26N6	661-27-51	C=C(C)C	InChI=15/4YSUBQYR2	Mass accu	TRUE					3412		1748	4244	
26	Actinobolol	C15H20N2	243-97-89	C1C@H	InChI=15/4PQVQBAW	Mass accu	TRUE	12772			17500	2760		3512		
27	Rifaximin	C45H51N3	88747-99	C1C@H	InChI=15/4NCRVPRKX	Mass accu	FALSE					4120	11830	1596		
28	ermetine	C29H40N2	481-18-1	CC(C@H)	InChI=15/4AUVVAXY	Mass accu	FALSE		4204				9176			9700
29	1,3-bis(4- C15H10N4	587-90-6	[O-][N+]	InChI=15/4JEZ20KXIX	Mass accu	TRUE		5772							5092	1760
30	Sulfabenz	C15H12N2	127-73-9	NC1ccc(cc1)	InChI=15/4PBCZLFBE	Mass accu	TRUE		1924			3964				

Intensity Retention Time Error_mDa Error_ppm QualifierFragments_mz_int_ext

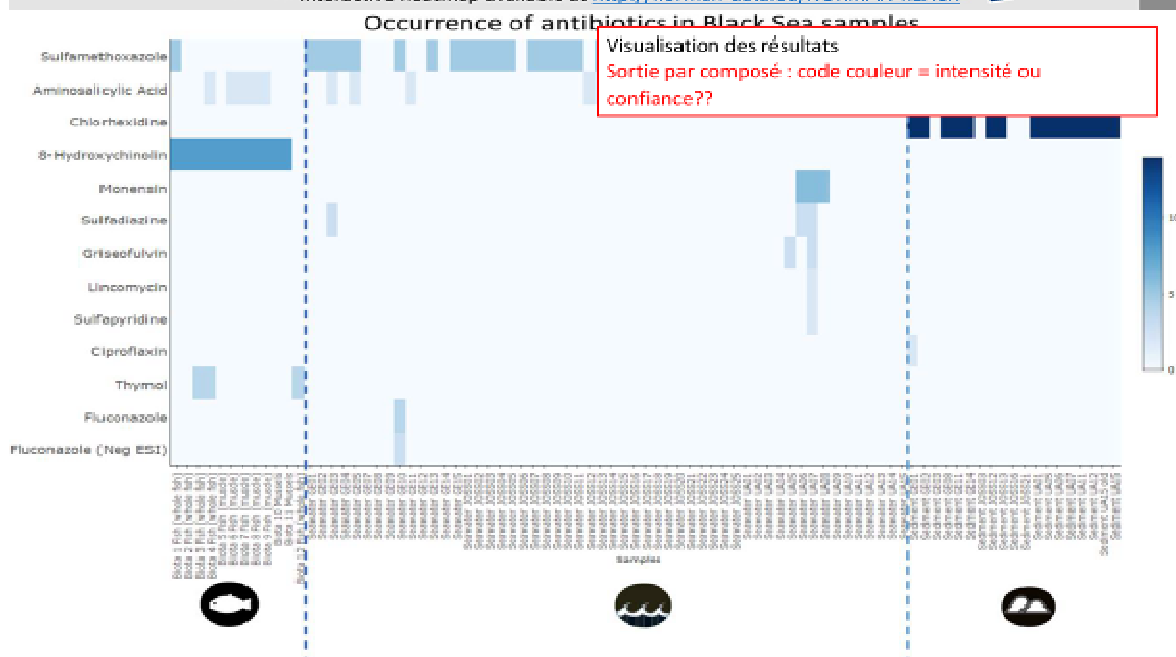
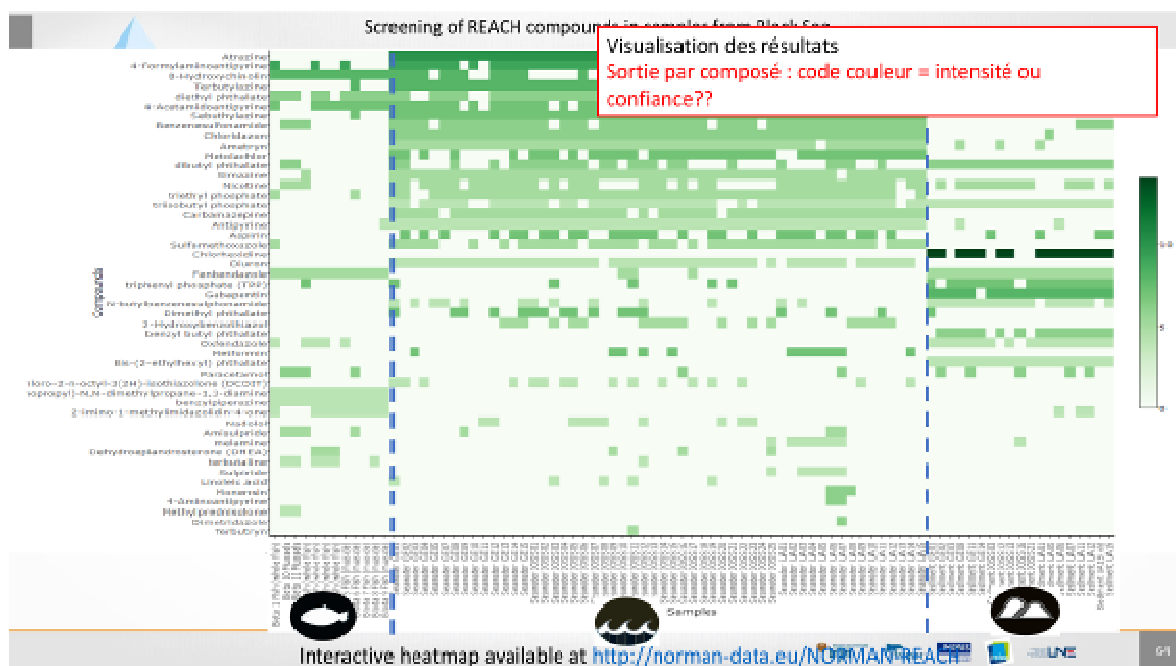
READY

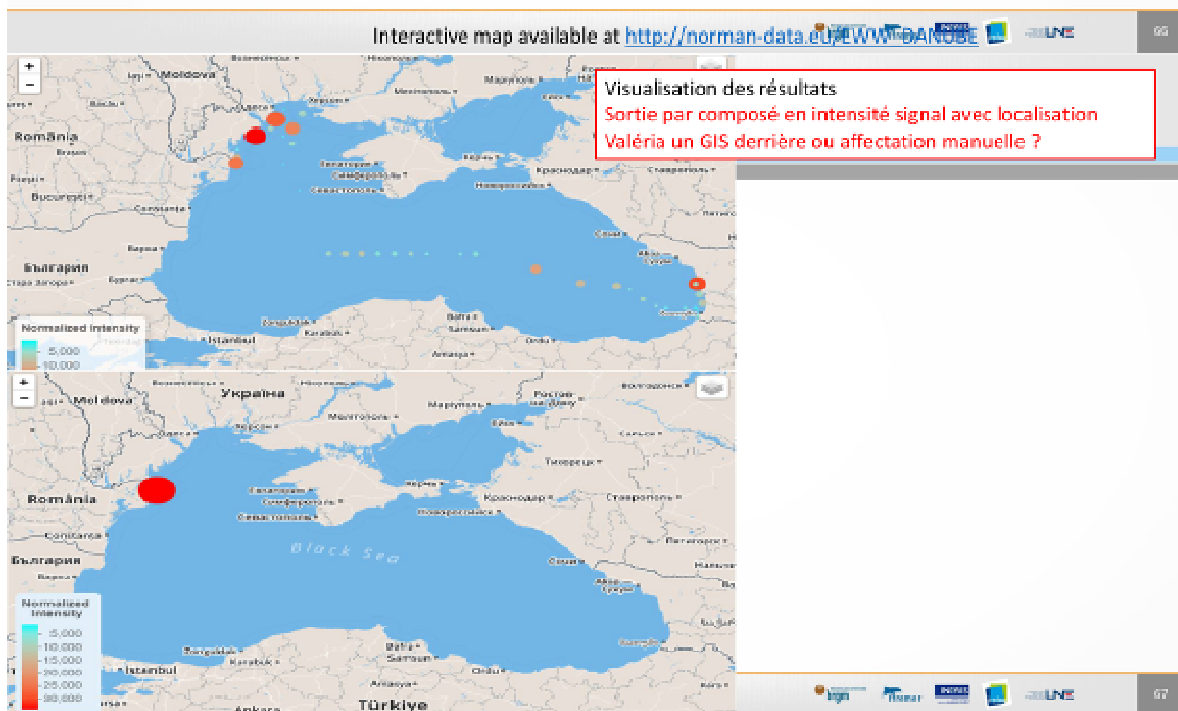
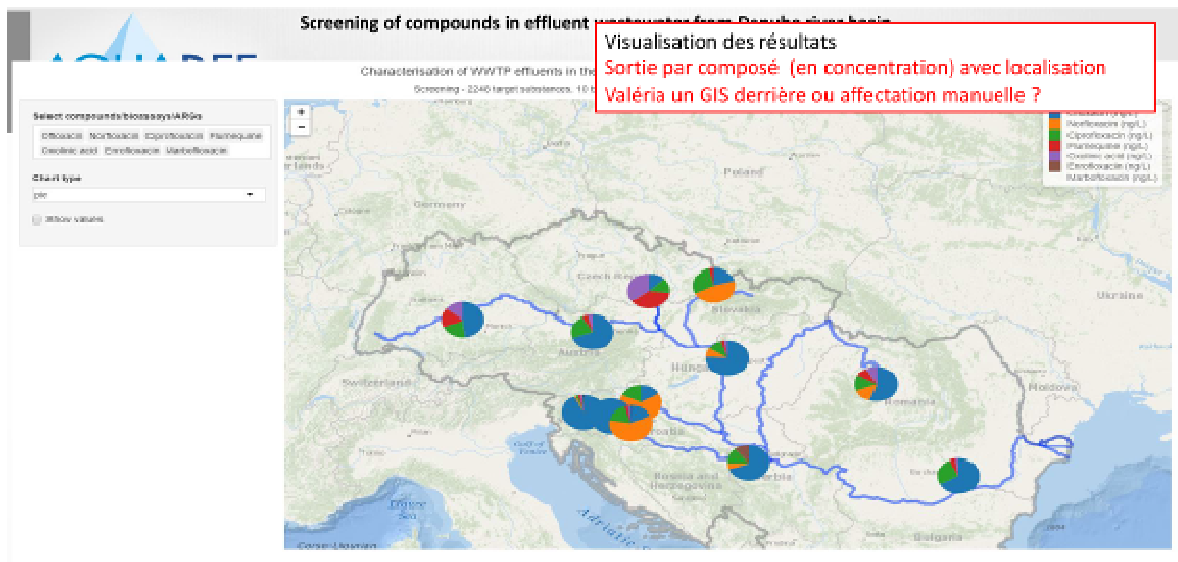
AQUAREF
Consortium
scientifique
de l'Environnement

ECM, PREMER,
PROM, BORD OF LIFE

Logo logos

Applications







Conseil scientifique
et technique

BRGM, IFREMER,
INRS, IRSTEA et UPA

Conclusions-Benefits of DSFP

Gestion effective des données sous forme d'un fichier non brut mxmml prétraité par le labo

Quid des blancs associés : possibilité de les bancariser pour laisser à l'utilisateur la possibilité de filtrer??

screening rapide de multitudes de couples substances/échantillons

- Attention risqué de comper des choux et des carottes (normalisation, effets matrices ...)

Update des listes suspect list avec le lien avec SUSDAT et suspect list

attention aux niveaux d'information très variables dans les listes.. Ex: molécules niveau 5 ne peuvent pas être utilisées

Possibilité de bénéficier des mises à jours fréquentes des informations sur les composés, sur les possibiités de retraitement

Semi- quanti : utilisation du critère d"homologie structurale...." pour PEC/PNEC OK mais sous forme de boîte noire...



Conseil scientifique
et technique

BRGM, IFREMER,
INRS, IRSTEA et UPA

Aspect technique DSFP

- Traitement automatisé des données approche ciblé/Suspect
mais pas de vrai NITS , comparaison d'empreintes etc...
- Gestion des réplcats ?
- pb transfert des fichiers Waters
- Non prise en compte des données GC-MS
- Qualification des données mises sur la base ?
- Traitement des données FR indépendemment des autres est possible (Filtres echantillons)
- Homogénéité nomenclature DSFP / bancarisation (ex code stations en SANDRE /Wise , numéro échantillon unique commun avec Bancaristion Sandre)
- Bancarisation NTS /EIP ?
- Bonne prise en compte des métadonnées associées au échantillon



Conseil scientifique
et technique

BRGM, IFREMER,
INRS, IRSTEA et UPA

Aspect technique DSFP



ASPECT juridique DSFP

traitement des données FR par d'autres est possible !

Données open access est la règle du jeu : qu'imagine t'on y mettre ?

Mais DSFP en propre semble possible .. Cf Allemagne?

Perennité de la base ? Source de financement de la base, pas d'entité juridique , pérenne...

Stratégie UBA : obligation de développer une base en propre, base DSFP en "miroir"

pas hébergé en France ?



Présentation banque BRGM

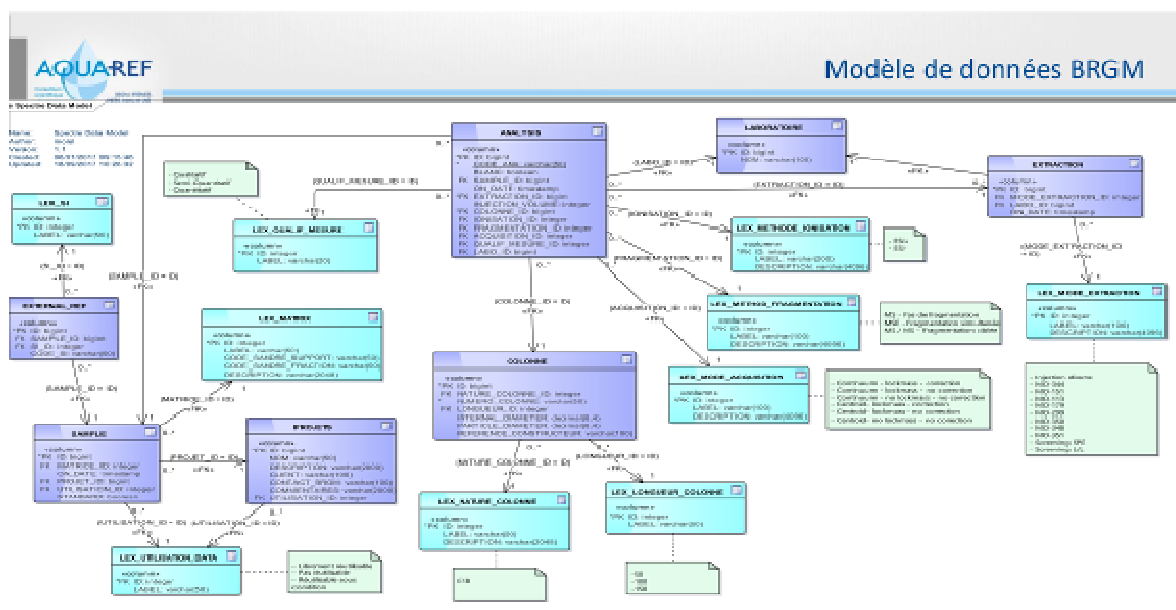
Créée pour un besoin interne de management de la donnée

Objectifs :

gérer le suivi analytique

permettre l'analyse retrospective et le regroupement d'échantillons issus de projets différents, selon date, localisation, matrice...etc





Spectrothèque BRGM

Analyses

Identifiant	Code Analyse	Statut	Date	Méthode toxicologie	Méthode fragmentation	Mode d'acquisition	Qualité
1	2014-12-17_STEP_Clemonet_044	Qualitatif	17/12/2014	MS/MS	MS/MS - Fragmentation simultanée	Controlé - lockdown - correction	Qual
2	2014-12-17_STEP_Clemonet_055	Qualitatif	17/12/2014	MS/MS	MS/MS - Fragmentation simultanée	Controlé - lockdown - correction	Qual
3	2014-12-17_STEP_Clemonet_056	Qualitatif	17/12/2014	MS/MS	MS/MS - Fragmentation simultanée	Controlé - lockdown - correction	Qual
4	2014-12-17_STEP_Clemonet_059	Qualitatif	17/12/2014	MS/MS	MS/MS - Fragmentation simultanée	Controlé - lockdown - correction	Qual
5	2014-12-17_STEP_Clemonet_059	Qualitatif	17/12/2014	MS/MS	MS/MS - Fragmentation simultanée	Controlé - lockdown - correction	Qual

Analyses

Identifiant	Code Analyse	Qualité de la mesure	Extraction	Volume injection	Laboratoire analyse	Recherche
1	2014-12-17_STEP_Clemonet_034	Qualitatif	1	2	BRGM	1
2	2014-12-17_STEP_Clemonet_095	Qualitatif	2	2	BRGM	2
3	2014-12-17_STEP_Clemonet_096	Qualitatif	3	2	BRGM	3
4	2014-12-17_STEP_Clemonet_096	Qualitatif	4	2	BRGM	4
5	2014-12-17_STEP_Clemonet_099	Qualitatif	5	2	BRGM	5
6	2014-12-17_STEP_Clemonet_010	Qualitatif	6	2	BRGM	6
7	2014-12-17_STEP_Clemonet_012	Qualitatif	7	2	BRGM	7
8	2014-12-17_STEP_Clemonet_013	Qualitatif	8	2	BRGM	8



Consortium scientifique
BRGM PREMIER

Information échantillon
Projet
Etc...

Créer ou éditer une analyse

ID 1

Code Analyse: 2014-12-17_STEP_Clermont_004

Blanc: ☐

On Date: 2014-12-17

Injection Volume: 2

Laboratoire: BRGM

Echantillon

Identifiant: 1

Date: 17/12/2014

Standard: false

Utilisation: Réutilisable sous condition

Matrice: EAU / Eau phase dissoute

Projet: 2014_APPLICATIONS_ENV

Réf externe:

[Sélectionner un échantillon](#)

Extraction:

Mode d'acquisition: Centroid - Isotopie

Méthode de fragmentation: MSE - Fragmentation

Colonne: 301

Méthode de ionisation: ESI+

Qualif Mesure: Qualité

[Annuler](#) [Sauvegarder](#)



Consortium scientifique
BRGM PREMIER

- Pas de traitement de la donnée : données brutes et besoin d'un labo expert pour traiter la donnée
- du coup on fait ce que l'on veut (comparaison d'empreintes, NTS...)
- pas de pb de volumes de données à stocker
- Orientés pour l'instant QA/QC "interne"
- Pas Volet interface GIS
- pas Interface accès extérieur prévu
- Peu descriptif d'échantillons(lien LIMS interne)

Annexe 2: La gestion des données : contexte et exigences

En vertu de la Directive **Loi Cadre sur l'Eau DCE**, les États membres doivent mener de vastes consultations auprès du public et des parties concernées en vue de déterminer les problèmes, les solutions et leurs coûts qu'il convient d'inclure dans le plan de gestion de district hydrographique. Pour ce faire, une vaste consultation d'au moins six mois doit être menée sur les plans de gestion de district hydraulique au cours de leur élaboration, puis tous les six ans, lors de leur mise à jour.

La mobilisation et la participation du public sont des conditions préalables à la protection des eaux. Sans le soutien de la population, les mesures réglementaires sont vouées à l'échec. Les citoyens européens doivent jouer un rôle essentiel dans la mise en œuvre de la DCE en aidant les gouvernements à trouver un équilibre entre les questions sociales, environnementales et économiques à prendre en considération.

Le SIE Le Système d'Information sur l'Eau (SIE) est un dispositif créé par l'État pour le partage et la mise à disposition des données sur l'eau du secteur public. Le schéma national des données sur l'eau (**SNDE**) fixe les objectifs, le périmètre, les modalités de gouvernance du système d'information sur l'eau (SIE) et décrit ses dispositifs techniques (de recueil, conservation et diffusion des données et des indicateurs). Il précise comment ces dispositifs sont mis en œuvre, comment les méthodologies et le référentiel des données et des services sont élaborés, et comment les données sont échangées avec d'autres systèmes d'information (<https://www.eaufrance.fr/le-systeme-dinformation-sur-leau-sie>)

Eau France est un service public d'information sur l'eau et les milieux aquatiques. Il a pour but de faciliter l'accès à l'information publique dans le domaine de l'eau en France. (<https://www.eaufrance.fr/propos-deaufrance>)

La **Directive INSPIRE**, élaborée par la Direction générale de l'environnement de la Commission européenne, vise à établir en Europe une infrastructure de données géographiques pour assurer l'interopérabilité entre bases de données et faciliter la diffusion, la disponibilité, l'utilisation et la réutilisation de l'information géographique en Europe. La Directive se structure en cinq parties correspondant aux composantes d'une infrastructure de données géographiques :

- Les métadonnées : elles sont la porte d'entrée de l'infrastructure puisqu'elles permettent de connaître les données et les services disponibles ainsi que leurs utilisations possibles,
- Les données géographiques : elles doivent être disponibles dans des formats et des structures harmonisés afin d'en faciliter l'utilisation par tous,
- Les services en ligne : toutes les données et métadonnées doivent être accessibles via Internet, vecteur privilégié d'échange,
- Le partage entre autorités publiques : les principes d'échange, de tarification et les conditions d'utilisation doivent faciliter l'accès aux données et aux services en ligne,
- Les mécanismes de coordination et de suivi de la directive : il s'avère nécessaire de mettre en place des structures de coordination tant des contributeurs que des utilisateurs (http://cnig.gouv.fr/?page_id=8991).