

Virtual Research Environment

Plateforme de travail collaborative

Plan

- Définition d'un VRE
- Exemples de VRE
- Use Case – Définition des besoins à l'Ifremer
 - Espace de travail et sélection de données
 - Communication
 - Traitements
 - Workflow
 - Labs
- Challenges
 - Fonctionnels
 - Techniques

Définition d'un VRE

VRE (Virtual Research Environment) ou CWE (Collaborative Working Environment)

- Environnement web permettant aux chercheurs de travailler en collaboration
 - Hébergement de fichiers/ressources
 - Outils spécifiques à la discipline associé (Jupyter Notebook, ...)
 - Outils de publication, de présentations, de redactions ...

Exemple de VRE

- BlueBridge

The screenshot shows the BlueBridge Gateway website interface. At the top, there is a navigation bar with icons for home, folders, documents, and email, along with a search box for the News Feed. The user is logged in as 'Gael Leblan'. The main header features the 'BlueBRIDGE Gateway' logo and the text 'Powered by D4SCIENCE INFRASTRUCTURE'. Below this is a dark blue navigation menu with options: 'Welcome', 'Explore Virtual Research Environments' (selected), 'Catalogue', and 'About'. The main content area is titled 'Explore Virtual Research Environments' and features a section for 'gCubeApps', described as a series of free-to-use applications for data analysis. Below this, there are six VRE cards, each with an icon, a title, and buttons for 'Enter this VRE' and 'Info'. The cards are: 1. AnalyticsLab (cloud icon with data terms), 2. BiOnym (book icon with 'Az' and a fish), 3. BiodiversityLab (Erlenmeyer flask with a fish and jellyfish), 4. RPrototypingLab (cloud icon with 'Rproto Lab'), 5. ScalableDataMining (public icon, binary code, and server rack), 6. TabularDataLab (public icon, cloud, and data table). A vertical sidebar on the right edge contains the text 'Report an issue'. At the bottom right, there is a Windows system message: 'Activer Windows. Accédez aux paramètres pour activer Windows.'

Exemple de VRE

- BlueBridge

The screenshot displays the BiodiversityLab VRE interface. At the top, there is a navigation bar with icons for home, folders, documents, and mail, along with a search box for the News Feed. The user's name, Gael Leblan, is visible in the top right corner. Below the navigation bar is a blue header with the following menu items: BiodiversityLab, Members, Species Data, Species Maps, Data Miner, Stat Alg Importer, Data Catalogue, R Studio, and Search.

The main content area is divided into several sections:

- Statistics:** "Your Stats in BiodiversityLab" showing activity and GOT (Goals On Track) with icons for a person, a thumbs up, and a speech bubble.
- Share updates:** A text input field for sharing updates, a dropdown menu for sharing options (currently set to "BiodiversityLab + Notification to members"), and a "Share" button.
- News feed:** A list of recent updates. The first update is from Gianpaolo Coro, dated March 28, 3:40 PM, with 2 likes. The second update is from Gianpaolo Coro, dated February 14, 3:40 PM, with 1 like. The third update is from Anton Ellenbroek, dated February 15, 10:22 AM.
- Top Topics:** "No Topics found in News Feed".
- Recent Documents:** A list of documents including "Schedule.txt", "results.csv", "Mola mola (scientific name)...", "faoceanarea_T20.csv", and "UPMC_Coro_iMarine2015_V2.0...".
- Authorisation Options:** A section for "Personal Token" with an "About Personal Token" box explaining its use for programmatic interaction.
- About:** A section with a logo of a beaker containing a fish and a jellyfish, and text describing the VRE's purpose: "The BiodiversityLab is a VRE designed to provide a collection of applications that allow scholars to perform complete experiments about single individuals or groups of marine species. The VRE allows to: inspect species maps; produce a species distribution map by means of either an expert system (AquaMaps) or a machine learning model (e.g. Neural Networks);".
- VRE Managers and Groups:** A section with "View Managers" and "Data-Manager" buttons.

On the right side of the interface, there is a vertical button labeled "Report an Issue".

Exemple de VRE

- BlueBridge

The screenshot displays the BlueBridge VRE interface. At the top, there is a navigation bar with icons for home, folders, documents, and mail, along with a search box for the News Feed. The user's name, Gael Leblan, is visible in the top right corner. Below this is a blue navigation bar with various menu items: BiodiversityLab, Members, Species Data, Species Maps, Data Miner (selected), Stat Alg Importer, Data Catalogue, R Studio, Search, and Giant Squid History. The main interface is titled 'DataMiner' and features a 'go back' button, 'Access to the Data Space', 'Execute an Experiment', 'Check the Computations', and 'Help' options. The central workspace is divided into two main sections: 'Operators' on the left and 'Parameters' on the right. The 'Operators' section lists several categories: BAYESIAN METHODS (7), CHARTS (4), CLIMATE (3), CORRELATION ANALYSIS (1), DATA CLUSTERING (4), DATABASES (4), and FIGIS SPREAD (3). The 'CHARTS' category is expanded, showing 'Generic Charts' as the selected operator. The 'Parameters' section for 'Generic Charts' includes a description: 'An algorithm producing generic charts of attributes vs. quantities. Charts are displayed per quantity column. Histograms, Scattering and Radar charts are produced for the top ten quantities. A gaussian distribution reports overall statistics for the quantities.' The parameters are: 'InputTable' (with a 'Select Data Set' button and a description of the input table format), 'TopElementsNumber' (set to 10), 'Attributes' (with an 'Integer Value' input field), and 'Quantities' (with a 'Select table from parameter InputTable' button). A 'Start Computation' button is located at the bottom of the parameter section.

Use Case – Définition des besoins à l'Ifremer

- Plusieurs projets Ifremer souhaitent mettre en place un VRE



- SeaDataNet Water Column Analysis => Analyse de la colonne d'eau (Temp et Salinité)
- SeaDataNet Biology quality assessment => Validation biologique statistique



- Emodnet Ingestion => Outils de conversion de format propriétaire



- Emodnet Chemistry => Analyse de la colonne d'eau sur des paramètres bio-geo-chimique
- Marinet 2 => Infrastructure de test en bassin

Labs

- Existe plusieurs environnements différents
 - Contexte différents
 - Outils en partie différents
 - Pratiques potentiellement différentes
- ⇒ Mise en place de « Labs », VRE avec ergonomie commune mais fonctionnalités différentes

Espace de travail et sélection de données

- Permettre à l'utilisateur de gérer ses fichiers de données via :
 - Possibilité de les consulter (ouverture de l'outil associé au type de fichier)
 - Possibilité de les partager avec d'autres personnes du VRE
 - Possibilité de sélectionner les données pour un usage dans d'autres aspect du VRE

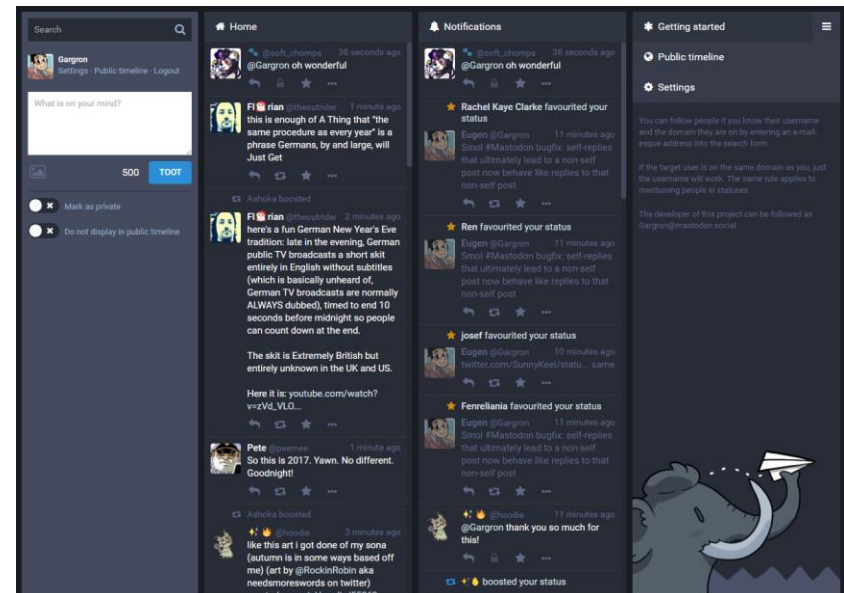
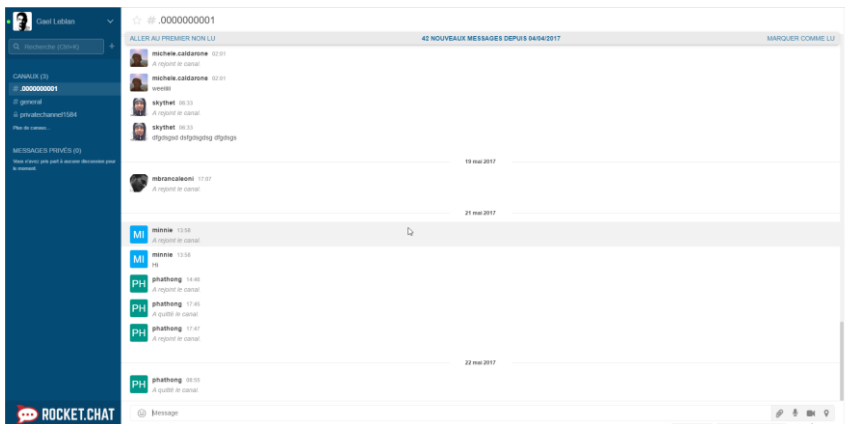
Espace de travail et sélection de données

- Permettre à l'utilisateur de sélectionner des données sur une carte
 - Outils de sélection différents suivant le lab (Sextant, Coriolis, ...)

The screenshot displays the Sextant web application interface. At the top, there is a header with the Sextant logo (a sextant) and the text "Sextant Infrastructure de données géographiques marines et littorales". To the right of the header, there are language options "EN FR" and the Ifremer logo. Below the header, there is a navigation bar with three buttons: "Présentation", "Services OGC", and "Ressources documentaires". A search bar with the text "Recherche" is also present. The main content area is divided into two sections. On the left, there is a sidebar menu with the following items: "Fond de carte : Sextant", "Activités humaines", "Campagnes océanographiques", and "Sampling operations". Below the sidebar, there are buttons for "LÉGENDES", "ORDRE", and "SOURCES". The main area on the right shows a map of the Atlantic Ocean with numerous data points plotted. The map is overlaid on a satellite-style background. At the bottom of the map, there is a scale bar indicating "2000 km" and a coordinate system indicator "WGS84".

Communication

- Communication au sein d'une équipe
 - Nécessité d'organiser des discussions au sein d'une équipe
 - Mais aussi discussions privés entre deux chercheurs
- => Communication avec Rocket.chat (Slack ou IRC-like) ou Mastodon (twitter-like)



Communication

- Possibilité de référence dans les conversations
 - Vers les données
 - Vers les traitements
 - Vers des workflow

=> Ouverture de l'outil associé

Traitements

- L'utilisateur doit pouvoir lancer des traitements et le VRE doit pouvoir lui en donner la possibilité
- Permet de s'affranchir des contraintes de la machine « client »
- Important de lui offrir les mêmes possibilités
- Attention aux outils qui peuvent permettre l'exécution de code arbitraire sur le serveur => Risque de code malicieux

Traitements

- Plusieurs outils peuvent être envisagés



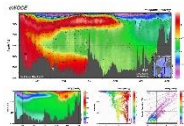
- Jupyter : écriture et exécution de python sur une machine serveur à partir d'un client web



- Beaker : proche de Jupyter avec plus de langage disponibles



- Globe : traitement de MNT (Modèle Numérique de Terrain)

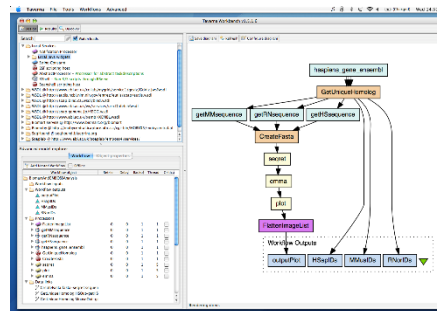


- ODV : Exploration, Analyse et Visualisation de données océanographiques

- De nombreux autres outils peuvent être envisagés, spécifiques au projet ou non ...

Workflow

- Certains traitements nécessitent d'être chaînés
 - Cela permet de créer une chaîne de traitement sur des données
 - Possibilité de lancer un traitement plusieurs fois d'affilé
 - Possibilité de rejouer un traitement



POC

- Premiers elements d'un POC

The image displays two screenshots of a Jupyter Notebook interface. The top screenshot shows a code cell with the following code:

```
In [2]: a = 15  
print(a)
```

The output is `15`. Below the code cell is a prompt: `Type Markdown and LaTeX: a^2` . The second code cell contains:

```
In [1]: b = 20  
print(b*2)
```

The output is `40`. The third code cell is empty: `In []:`. The bottom screenshot shows the same notebook with a sidebar on the left. The sidebar contains the following sections:

- MESSAGING
- FILE SYSTEM
- SELECTING DATAS
- CORIOLIS
- OCEANBROWSER (highlighted)
- TOOLS
- JUPYTER NOTEBOOK
 - MYNOTEBOOK (highlighted)
- ADD A NOTEBOOK
- ODV
- WORKFLOW

The main notebook area shows the same code and output as the top screenshot.

Challenges Fonctionnels

- Cette nouvelle façon de travailler entraîne de nombreuses problématiques fonctionnelles
 - Reflexion sur la manière de travailler de l'utilisateur
- Pour passer la résistance au changement :
 - Simplicité de prise en main
 - Environnement attractif
 - Garder une cohérence entre les outils (graphique ou fonctionnelle)

Challenges Techniques

- Isolation de la session de l'utilisateur
- Gestion de la donnée
 - Upload de donnée volumineuse
 - Localisation de ces données/traitement => rapidité
- Difficulté à faire travailler tout les outils ensemble
 - Obligation de mettre en place un niveau d'abstraction (Data Package, Process Package ?...)
 - Utilisation de standards (WPS, ...)

Questions