



Publication de codes sources à l'OFB

*Un aperçu des procédures et
outils existants*

Cellule Outils 24/10/2020

[1] Publication de code source, *open source*, « code ouvert »... de quoi parle-t-on ?

[2] Un inventaire (partiel) des pratiques à l'OFB

[3] Cas en externe – des exemples à suivre ?

[4] Des collaborations et des échanges, plutôt que des *préconisations*

[1] Publication de code source, *open source*, « code ouvert »... de quoi parle-t-on ?

Une philosophie, un principe de conception



La désignation **open source**, ou **code source ouvert**, s'applique aux logiciels (et s'étend maintenant aux œuvres de l'esprit) dont la licence respecte des critères précisément établis par l'Open Source Initiative, c'est-à-dire les possibilités de libre redistribution, d'accès au code source et de création de travaux dérivés. Mis à la disposition du grand public, ce code source est généralement le résultat d'une collaboration entre programmeurs.

Wikipedia d'après BTQ

Des licences de logiciels et de réutilisation

4 libertés fondamentales :

- utiliser le logiciel ;
- étudier le logiciel ;
- copier le logiciel ;
- modifier le logiciel et de redistribuer les versions modifiées.

d'après Open Source Initiative



Richard Stallman
« C'est plus vertueux »



Eric Steven Raymond
« Ça marche mieux. »

1983 : GNU

1985 : FSF

1991 : Linux

1998 : Open Source

38

Wikipedia recense

licences différentes,

différenciées par les notions de **compatibilité**, de **copyleft**, de **permissivité**.



Quid de la **licence ouverte** ?

<https://www.data.gouv.fr/fr/licences>



[1] Publication de code source, *open source*, « code ouvert »... de quoi parle-t-on ?

Pour quoi ?

Pour sécuriser
Pour partager et se faire connaître
Pour inciter et *populariser*
Pour s'approprier
Pour améliorer, pour s'améliorer
Pour contribuer et participer,
Pour intégrer les communautés

...

Développeurs, contributeurs, intégrateurs ou utilisateurs ?

Pour qui ?

Ingénieurs, développeurs
Scientifiques
Pilotes de projets
Architectes logiciels et données
Services infra et hébergement
Ceux qui viennent après (transfert)

...

En interne ou vers l'externe ?

Des conventions et associations
aux échelles internationales



Une certaine
attention politique



Des outils collaboratifs à monter sur diverses
plateformes locales ou mutualisées

[2] Un inventaire (partiel) des pratiques à l'OFB

Des rôles et des objectifs différents

	Besoins internes	Publications externes
Développeurs	Créer des fonctionnalités	Porter un projet publiquement Demander l'aide des communautés Ouvrir ses productions
Contributeurs	Améliorer et ajouter des fonctionnalités Bénéficier de l'existant	Ajouter ses propres fonctionnalités Ouvrir ses productions
Intégrateurs	Echanger du code entre membres d'un projet et/ou avec prestataire Documenter et rendre reproductible	Récupérer les fonctionnalités Montées de version Devenir contributeur
Utilisateurs	Récupérer les fonctionnalités Montées de version Suivre les évolutions, services spécifiques	Récupérer les fonctionnalités Montées de version

NB. Non exhaustif, frontières très mixtes et floues, pas de découpage « tel service ici ».

[2] Un inventaire (partiel) des pratiques à l'OFB

Quelques pratiques identifiées : plateformes publiques généralistes

DEVELOPPEURS ET CONTRIBUTEURS, VERS L'EXTERNE

Développeurs et contributeurs cherchant à valoriser leurs productions et à créer ou rejoindre une communauté (dite *en solo*)

Biométriciens et ingénieurs experts
Statisticiens, étudiants
Innovations (projets étudiants)

Thèses

<https://github.com/VajasPablo/catcheffortwb>

AdeHabitat



<https://github.com/ClementCalenge/adehabitat>

Aquadrone

<https://github.com/collombj/aquadrone-qgis>
<https://github.com/collombj/aquadrone-positionnement>



Outils R et Shiny

<https://github.com/CedricMondy>

ETL

<https://github.com/AlexandreBjLiccardi/wild7>

[2] Un inventaire (partiel) des pratiques à l'OFB

Quelques pratiques identifiées : plateformes publiques généralistes

DEVELOPPEURS ET CONTRIBUTEURS, VERS L'EXTERNE

**Par des groupes de développeurs, contributeurs et intégrateurs,
pour des projets ou des pôles RDI**

Le gitLab OFB en mode public

Rapportages

Pôle INSIDE et innovation

Datavisualisation

Réseaux

Le cas Hub'eau

SI fédérateurs, projets (ex. SINP-GINCO)

<https://github.com/INSIDE-information-systems>

<https://gitlab.afbiodiversite.fr/>

<https://github.com/AFB-dataviz>

<https://github.com/BRGM/hubeau>

<https://github.com/SINP-GINCO>



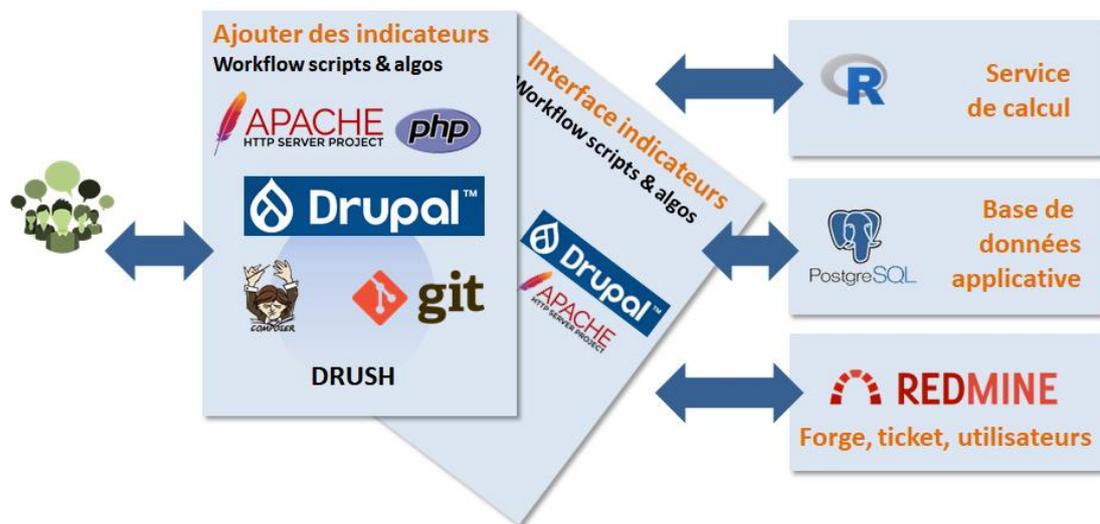
[2] Un inventaire (partiel) des pratiques à l'OFB

Quelques pratiques identifiées : sur des plateformes spécialisées

DEVELOPPEURS ET CONTRIBUTEURS, VERS L'EXTERNE

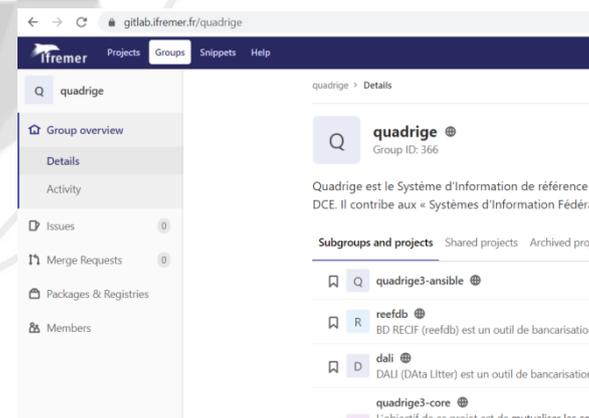
Par des groupes de développeurs, contributeurs et intégrateurs, pour des projets d'évaluation

Le cas du Système d'Evaluation de l'Etat des Eaux



Quid du contexte des SI fédérateurs ?

<https://gitlab.ifremer.fr/quadrige>



[2] Un inventaire (partiel) des pratiques à l'OFB

Quelques pratiques identifiées : sur des plateformes en interne

DEVELOPPEURS, CONTRIBUTEURS ET INTEGRATEURS, VERS L'INTERNE

Par des groupes de développeurs, contributeurs et intégrateurs, pour des projets

Le gitLab OFB en mode privé

Utilisation d'outils collaboratifs pour l'exécution de marché publics (*administration des données, bibliothèque ETL WILD*)

Utilisation d'outils collaboratifs par l'UMS Patinat et l'uDAM

Cap sur Jupyter (et les notebooks)



765

Dépôt de code et suivi de projet ?

projets sur la GED OFB

106

projets sur les forges

20

projets sur le gitLab

Et plus largement des outils

Maven



[3] Cas en externe – des exemples à suivre ?

BRGM

En externe : <https://github.com/BRGM>

Pour les logiciels *open source*, le code est accessible via les sites dédiés

<https://www.brgm.fr/production-scientifique/logiciels-scientifiques/logiciels-scientifiques>

En interne, GitLab : <https://gitlab.brgm.fr/>

Les **schémas** font l'objet de normes et sont communiqués via les SI référents ou l'OGC.



Depuis l'extérieur, un succès participatif modéré des différents projets – contributions essentiellement sur **ComPASS**, **razorback**.

CEREMA

En externe : <https://github.com/CEREMA>, terminologie spécifique des projets

Aussi, des projets plus rares via ADULLACT :

<https://gitlab.adullact.net/prodigeadmin/Prodige> et FRAMAGIT :

<https://framagit.org/cerema-dev/aistool>

La démarche *open source* est bien identifiée et centralisée dans les communications de l'établissement :

<https://www.cerema.fr/fr/actualites/diffusion-open-source-chaine-traitement-images-satellites-du>

Les **schémas** sont produits conjointement avec le CNIG et la COVADIS, dans un cadre INSPIRE (essentiellement *équipement*).



Depuis l'extérieur un nombre important de projets (> 50) et de participants, mais peu de reprise (et de **forks**) des codes sources.

[3] Cas en externe – des exemples à suivre ?

IGN

En externe : <https://github.com/IGNF>, des projets phares (plugin QGIS, geoportal)
Voir aussi la dynamique *lab*, avec <https://www.ign.fr/ignfab>. Les contributeurs se retrouvent aussi sur certains dépôts framagit.

La démarche *open source* est bien identifiée et centralisée dans les communications de l'établissement (licence CeCILL), avec un déploiement de l'offre à côté des logiciels commerciaux : <http://logiciels.ign.fr/>

Les **schémas** sont produits conjointement avec le CNIG et la COVADIS, dans un cadre INSPIRE.



Depuis l'extérieur un nombre important de projets (> 50) et de participants, quelques projets phares et faisant l'objet de reprises (*fork*).

Universités



Des situations variées. De manière générale, on retrouve souvent la structure :

- gitLab en propre, plus au moins fermés, pour les étudiants et personnels
- framagit et gitHub, pour les étudiants en autonome et pour des projets ou des filières très clairement identifiées.

Exemples de Lille, Lyon, Montpellier...

[3] Cas en externe – des exemples à suivre ?

Partenaires : PNx

Volonté de centralisation des outils de diffusion et de communication sur la base *framsoft*, et les codes sources sous : <https://github.com/PnX-SI> pour les projets principaux, les parcs nationaux disposant sinon de gitHub autonomes.

Certains projets font l'objet de dépôts individuels (<https://github.com/GeotrekCE>) ou via le prestataire (Makina Corpus) pour des raisons de gouvernance.



**Au total, plus de 50 dépôts particulièrement actifs.
Depuis l'extérieur, un effort d'animation et une coordination visibles.**

NB. Bien des pistes n'ont pas été explorées ici ! Comme le github commun aux Ministères de la Transition écologique & de la Cohésion des territoires <https://github.com/MTES-MCT>, le gitLab IRSTEA <https://gitlab.irstea.fr/> (quid INRAE ?), IFREMER (<https://gitlab.ifremer.fr>)...

[3] Cas en externe – des exemples à suivre ?

Etalab – licences, ressources et recommandations



Des recommandations quant au licences :

<https://www.data.gouv.fr/fr/licences>

Des productions en appui aux EIG et AMI

Le github d'Etalab, et ses outils en ligne dont le guide

<https://github.com/etalab>

<https://github.com/betagouv> (contributions possibles)

<https://github.com/etalab/etalab/blob/master/logiciels-libres.md>

<https://github.com/etalab/ouverture-des-codes-sources-publics>

Référencer les schéma de données publiques

<https://schema.data.gouv.fr/>

Des retours d'expérience

<https://entrepreneur-interet-general.etalab.gouv.fr/blog/2018/11/14/administrations-eig-et-logiciel-libre.html>

Les ressources de la DINUM

<https://github.com/disic/politique-de-contribution-open-source/>

La communauté des contributeurs au logiciel libre dans l'administration (*Blue Hats*) :

<https://www.numerique.gouv.fr/actualites/la-communaute-blue-hats-hackers-dinteret-general-est-lancee-rejoignez-nous/>

Une plateforme de référencement des codes sources du secteur public

<https://code.etalab.gouv.fr/>

Licences permissives	identifiant SPDX
Apache License 2.0	Apache-2.0
BSD 2-Clause "Simplified" License	BSD-2-Clause
BSD 3-Clause "New" or "Revised" License	BSD-3-Clause
CeCILL-B Free Software License Agreement	CECILL-B
MIT License	MIT

Licences avec obligation de réciprocité	identifiant SPDX
CeCILL Free Software License Agreement v2.1	CECILL-2.1
CeCILL-C Free Software License Agreement	CECILL-C
GNU General Public License v3.0 or later	GPL-3.0-or-later
GNU Lesser General Public License v3.0 or later	LGPL-3.0-or-later
GNU Affero General Public License v3.0 or later	AGPL-3.0-or-later
Mozilla Public License 2.0	MPL-2.0

[3] Cas en externe – des exemples à suivre ?

Mission *Politique publique de la donnée (dite Bothorel)*

<https://www.mission-open-data.fr/>, consultation du 8 octobre au 9 novembre 2020.

La synthèse est aujourd'hui essentiellement constituée de 10 constats, ainsi que des réactions issues de la consultation.

Il y a urgence à agir car les technologies progressent plus vite que le droit

Le manque de doctrine claire et partagée limite la cohérence de la politique publique de la donnée

Le manque de culture et de compétences limite la libération du potentiel de la donnée et des codes sources

L'ouverture est perçue comme un danger par les acteurs publics, à tort ou à raison

Le cadre juridique, complexe, foisonnant et soumis à des interprétations divergentes, est difficile à maîtriser

Il manque un portage managérial et une formation des cadres dirigeants aux enjeux de la donnée

L'offre de données et de codes sources ouverts manque de lisibilité

Le portage politique de la politique de la donnée est insuffisant

Les ressources des acteurs publics sont contraintes, alors que l'ouverture demande des investissements et des moyens humains

Les acteurs publics ont des difficultés pour se doter en compétences dans la donnée et les codes sources

Rapport d'étape : https://www.mission-open-data.fr/uploads/decidim/attachment/file/2/Politique_publice_donn%C3%A9e_Rapport_%C3%A9tape_VF.pdf

[4] Des collaborations et des échanges, plutôt que des préconisations

Quelles obligations et quelles sollicitations réelles ?

Vous a-t-on déjà demandé accès à un code ?

Si oui, une administration, un acteur privé ou un particulier ?

A l'uDAM en 2020 : projet IA pour la Police de l'Environnement (administration, pour recherche et développement), documentation complémentaire SEEE (prestataires privés).

Des conditions (supplémentaires) pour la publication

L'absence d'intérêt commercial, ou de contrat en cours. L'OFB est susceptible d'avoir recours à des prestataires externes pour développer ses logiciels. Les conditions de publications peuvent être tenue par contrat durant la phase de développement.

La preuve de l'innocuité de la publication du code, ou à défaut une expertise permettant de garantir que les codes libérés **ne peuvent pas nuire à l'environnement** (voir les exemples précédents sur les espèces menacées et l'activité de Police). De la même manière, la mobilisation de données personnelles ne doit pas apparaître ou faire l'objet de rétroingénierie.

La maturité des codes publiés. Certains processus sont longs et complexes à interpréter, comme la mise au point d'algorithmes d'évaluation de l'état de l'environnement. Un accompagnement, une documentation et une certaine ergonomie doivent permettre la publication de code réutilisables et lisibles.

Exemple : IA pour la Police de l'Environnement

[4] Des collaborations et des échanges, plutôt que des *préconisations*

Publication de codes sources : des outils de *travail* plus que de *contrainte*

Chaque projet a sa gouvernance et ses règles en propre !

Mis à disposition par la DSI de l'OFB : GitLab public (en lecture), ou privé.
Voir avec pascal.perrier@ofb.gouv.fr

Se rapprocher sinon des différentes instances (réseaux, projets, SI fédérateurs, partenaires...).

Ne pas oublier qu'un code hébergé sur un gitLab « de prestataire » doit être récupéré et republié si possible !

Objectifs communs : permettre un maximum d'accès et de mutualisation, éviter la multiplication des outils

