



**Évaluation de la valorisation
par l'Open Source :**
enseignements tirés de l'étude
réalisée pour le CNES



Inno³ est un cabinet de conseil en **innovation ouverte**, indépendant et spécialiste de l'**Open Data** et de l'**Open Source**.

Nous nous impliquons auprès de **grands groupes**, **d'instituts de recherche**, de **jeunes pousses**, de **collectivités locales** et de collectifs divers pour construire à leurs côtés des **formes novatrices** de collaboration et d'ouverture.

Nous définissons ensemble des **stratégies ambitieuses** pour en tirer parti ainsi que des **politiques pragmatiques** qui permettent de les mettre en œuvre.

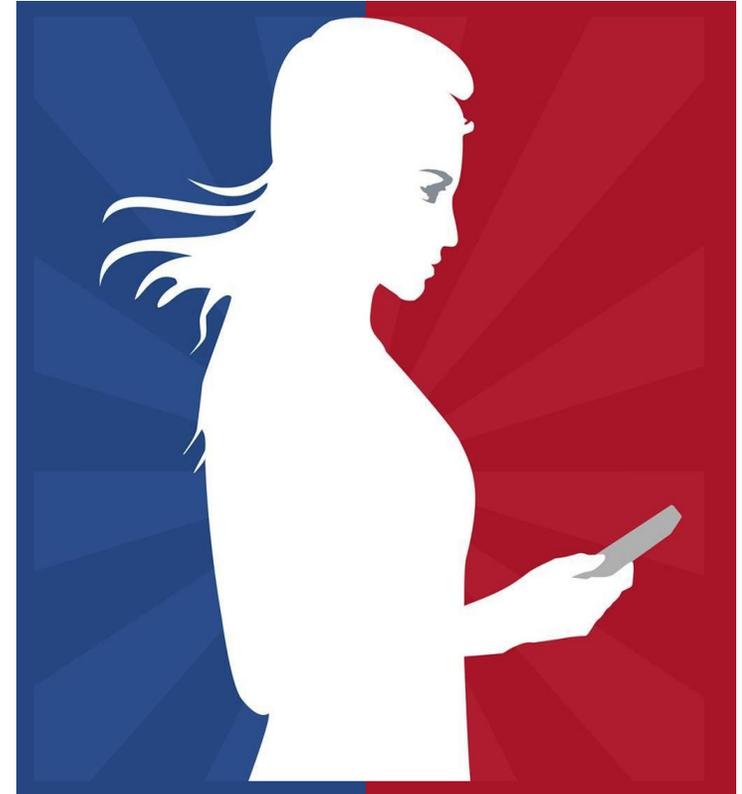


■ Contexte juridique : La Loi pour une République Numérique

LOI n° 2016-1321 du 7 octobre 2016 pour une République numérique

- Instaure l'ouverture par défaut des codes sources des logiciels des administrations (Art. 2) [entrée en vigueur le 7 oct. 2018]
- Introduit un « encouragement » au logiciel libre (Art. 16) :

Les administrations doivent encourager « l'utilisation des logiciels libres et des formats ouverts lors du développement, de l'achat ou de l'utilisation, de tout ou partie, de ces systèmes d'information. »



Dans son référé du 11 mars 2019, la Cour des Comptes a invité Météo France, l'IGN et le CEREMA à se saisir de cette opportunité pour faire évoluer leurs modèles, leurs pratiques, voire leurs rôles respectifs.

■ Contexte juridique : Autres éléments



- Cette tendance a été confirmée lors de la révision de la **directive relative à la réutilisation des données publiques** (dite Directive PSI) adoptée le 20 juin 2019
- Vise à renforcer le marché unique de la donnée – dans le prolongement du règlement sur la libre circulation des données – en **réduisant les barrières à la réutilisation** des données ouvertes.

■ Contexte général d'ouverture : Science ouverte et politiques publiques



- **Politique publique européenne**
Open Science avec Horizon 2020 et renforcée notamment en 2016 avec le rapport ***Open Science, Open Innovation, Open to the World.***
- En France, juillet 2018 lancement du Comité pour la Science Ouverte (CoSO) et du Plan National pour la science ouverte.



■ Enjeux : un changement en termes de valeur



- La notion de **valeur** (notion économique) est beaucoup plus subjective que celle de **propriété** (notion juridique).
- La valeur pour un centre de recherche est une combinaison de la valeur générée par le technologie et la valeur captée par le centre de recherche. L'Open Source et l'Open Data démultiplient la première et réduisent partiellement la seconde.
- **Ouvrir ne veut pas dire abandonner**
Les stratégie nouvelles doivent être anticipées en amont des typologies d'usage et des modalités associées. Elles doivent aussi clairement identifier l'emprise de chacun sur chaque ressource.
- **Combiner, articuler et ne pas opposer**
La stratégie repose sur la combinaison de différents actifs immatériels afin d'être optimisée.
- **Ne pas confondre posséder et contrôler**
La place prise dans la gouvernance d'une ressource prime sur la question de la propriété. L'adoption d'un projet par les tiers repose sur la capacité à partager la gouvernance de la ressource.



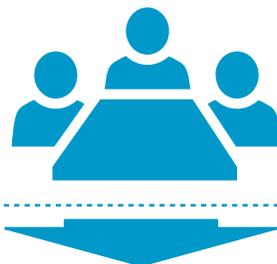
Reposant sur un nombre important d'acteur, les modèles ouverts impliquent de :

- trouver un juste milieu entre **anticipation** et **lâcher-prise** ;
 - Repenser les outils juridiques afin de combiner **autorité** et **agilité**
-
- **Assurer le développement maximal de la ressource : permettre à tous de contribuer et d'exploiter la ressource ;**
 - **Adapter le cadre de collaboration à ses propres ambitions**
 - **Un socle contractuel fort**
 - **Un socle interopérable avec d'autres approches**
 - **Avoir une vigilance particulière quant au respect des contraintes collectives.**

■ Méthodologie



Travaux antérieurs
Groupe de travail EvalAIO



etalab^{gouv.fr}

Principes transverses de valorisation Logiciel dans le contexte du CNES



université
PARIS-SACLAY





Création et captation de valeur



■ Qu'est-ce qui a de la valeur ?



- Valeur financière
 - Pas liée à un objectif de maximisation systématique
- **Autres aspects liés aux raisons d'être des établissements**
 - Diffusion de la connaissance
 - Progression de la recherche
 - Dynamisation économique d'un secteur
 - ...

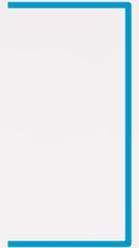
■ De la valeur pour qui ?



- **Périmètre interne / élargi**
 - À l'échelle d'un projet
 - À l'échelle de l'établissement
 - À l'échelle d'un groupe d'établissements...
- **Périmètres externes**
 - Captation de la valeur par des entités tierces :
 - Entreprises d'un secteur donné
 - Établissements d'enseignement/recherche
 - Public indifférencié (portée sociétale)
 - Pondération géopolitique de la captation de valeur



Critères de valorisation



■ Valeur corrélacionnelle



Des facteurs intrinsèquement liés à l'Open Source, dont l'impact est indéniablement attesté bien que difficilement mesurable.

- Innovation et agilité (time to market, etc.)
- Transparence (reproductibilité, etc.)
- Qualité
- Mutualisation interne

■ Valeur d'usage et diffusion



- Valeur par l'usage : élément fondamental de la valorisation de l'Open Source
- Libre diffusion : atout fondamental de l'Open Source, mais complexifie la mesure
- Mesure : souvent un frein de la diffusion.
- Des mesures complexes à interpréter :
 - Ordres de grandeur / classement relatif
 - Tendances

■ Évaluation : estimations quantitatives



- Téléchargements
 - Biais maximisant : téléchargement! = utilisation effective
 - Biais minimisant : multiplicité des canaux de distribution
- Pour les applications :
 - Télémétrie (Problème de confidentialité)
- Pour les bibliothèques :
 - Nombre de paquets dépendants

fs-extra

8.1.0 • Public • Published 4 months ago

Readme

3 Dependencies

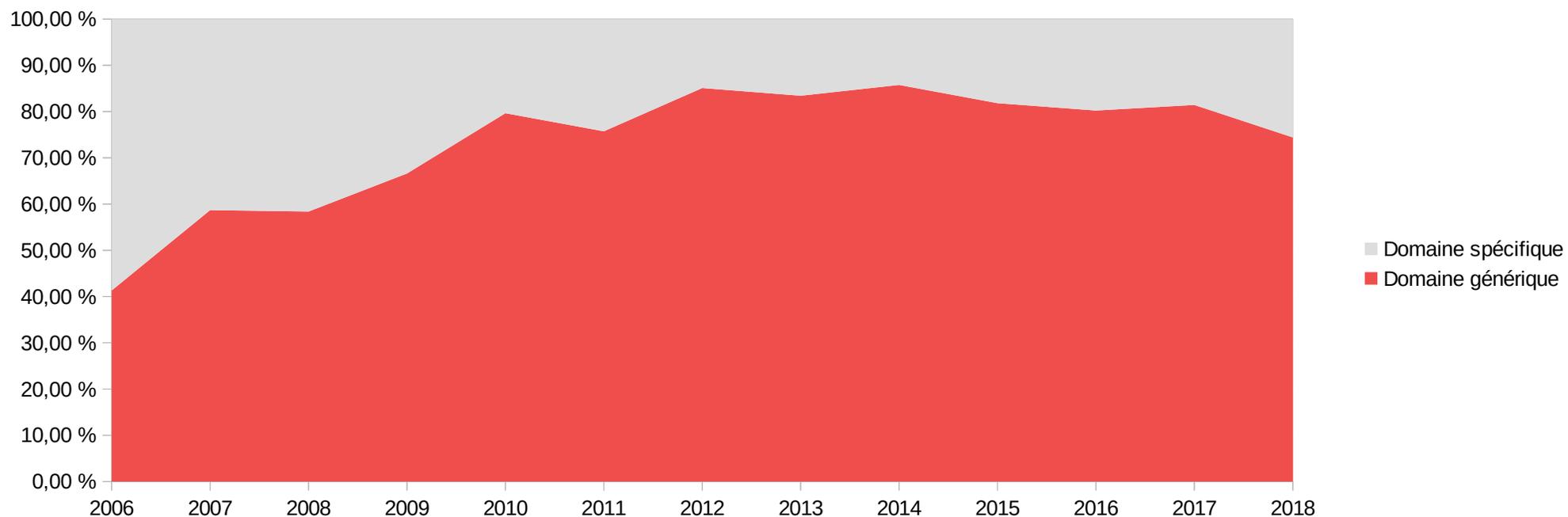
28 870 Dependents

■ Évaluation : qualification des usages



- Objectif : qualifier la population utilisatrice pour déterminer les aspects de création de valeur :
 - Distinction par les types d'utilisateurs
 - Distinction par localisation
- Moyens
 - Déclaratifs :
 - Questionnaires
 - Enquêtes
 - Analytiques :
 - IP des téléchargements
 - NDD pour les échanges sur les canaux de discussions

■ Évaluation : aspects pratiques



Domaines d'origine des messages sur les listes de discussion OTB

■ Évaluation : exemples



Exemple extrait de l'analyse d'OTB

Région	Grand groupe	Non profit	PME	Secteur public	Université / Recherche	Total
France	3		4	9	17	33
Europe	4		10		18	32
Amérique du Nord		2	9	1	8	20
Asie			2		4	6
Global			1		1	2
Afrique			1		3	4
Océanie			1		1	2
Europe hors UE			1			1

■ Développement de l'activité économique



- Mesures dérivées de celles des usages mentionnée
 - Restriction à un type d'entité
 - Interprétation par localisation
- Affinement par des métriques complémentaires spécifiques théoriquement possible
 - Chiffres d'affaire générés, emplois associés, etc.
 - En pratique, complexité d'obtention et de communication

■ Soutien à la recherche et rayonnement scientifique



- Mesures dérivées de celles des usages
 - Restriction à un type d'entité
 - Interprétation par localisation
- Mesures spécifiques :
 - Publications scientifiques sur le logiciel lui-même
 - Publications scientifiques citant le logiciel
 - La pratique de la citation des logiciels encore dans une phase d'évolution
- Contexte d'évolution des mesures et de la citation en recherche:
 - Citation publi > Données > Logiciels ? (tracabilité de la démarche de recherche et reproductibilité)
 - Mise en œuvre de plan de gestion de données (DMP) : ensemble de critères pouvant servir à l'évaluation / Mesure
 - Vers un SMP (Projet PRESOFT?)
 - Standards et normalisation des outils DOI (*Digital Object Identifiers*) accompagnées de « bonnes pratiques » open science : **FAIR DATA**, etc.
 - Initiatives spécifiques : « the FORCE11 Software Citation Implementation WG »

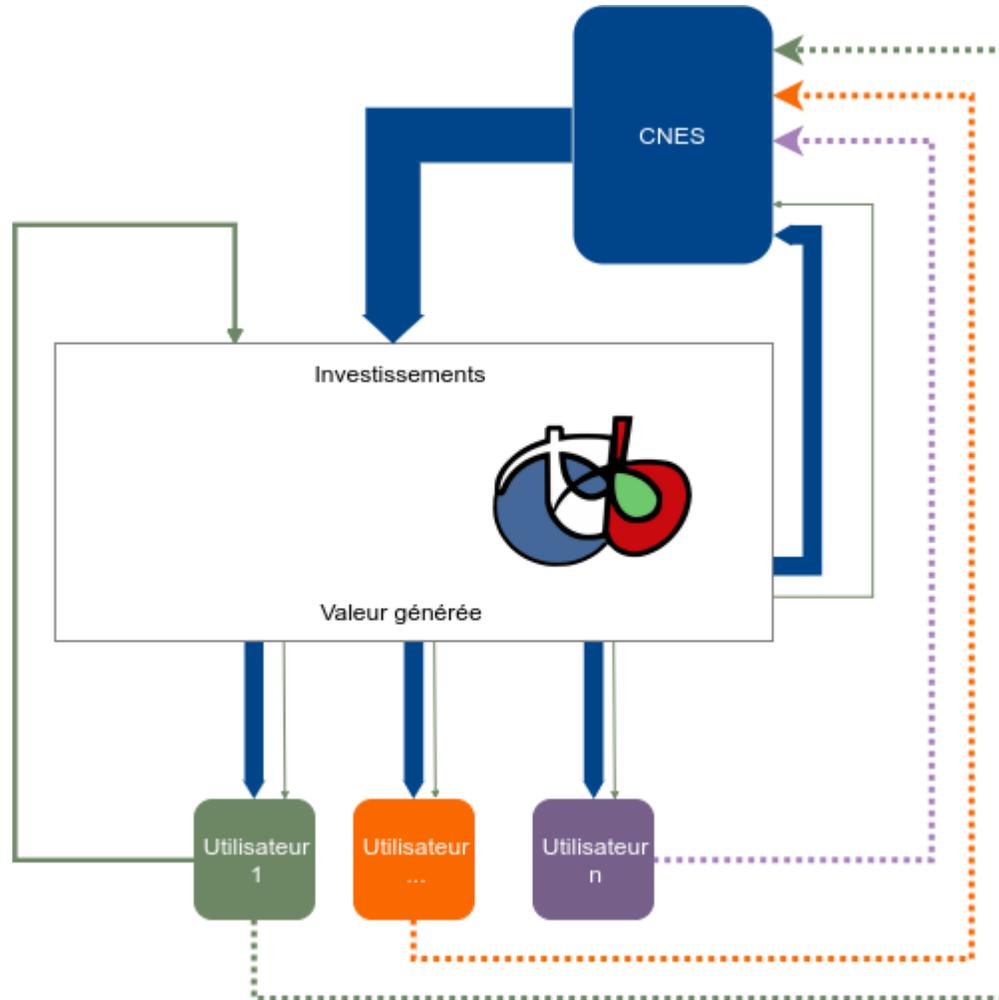


Cite US!

If you use scikit-multilearn in your research and publish it, please consider citing us, it will help us get funding for making the library better. The paper is available on [arXiv](#), to cite it try the Bibtex code on the right.

```
@ARTICLE{2017arXiv170201460S,  
  author = {{Szyma{\n}ski}, P. and {Kajdanowicz}, T.},  
  title = "{A scikit-based Python environment for performing multi-label classification}",  
  journal = {ArXiv e-prints},  
  archivePrefix = "arXiv",  
  eprint = {1702.01460},  
  primaryClass = "cs.LG",  
  keywords = {Computer Science - Learning, Computer Science - Mathematical Software},  
  year = 2017,  
  month = feb,  
}
```

■ Mutualisation externe



■ Mutualisation externe



- **Mesures possibles sur :**

- les volumes de contributions
- Les identités des contributeurs

- **Types de contributions :**

- Code
- Hors code (documentation, traduction, QA, communication, etc.)
- Services SaaS gratuits

- **Contributions de l'entité à des projets tiers**

- Meilleure diffusion de l'effort investi
- Possible dans un cadre de logiciel propriétaire mais plus complexe

■ Métriques de contributions de code



- **Git, standard de fait pour les développements**
 - Fournit de nombreuses informations exploitables
 - Et exploitées par différentes applications spécialisées
 - Gitstats
 - Façade OSS
 - Grimoirelab
 - Projet global CHAOSS (Linux Foundation)
- **Question du choix de micrométriques :**
 - Nb & type de contributeurs / entités contributrices entités de des contributeurs
 - Volume de contributions (lignes ajoutées ?)
- **Limites du traitement automatique :**
 - informations à interpréter avec les mainteneurs du projet.

Métriques de contributions (code et +)



Overview

*

project: "cnes" Actions

Welcome!

Click on cells, bars, slices, drag & drop on widgets or set a time range on the top right corner to filter data. Filters appear above as small stickers. [More](#)

Top Submitters (Issues & PRs)

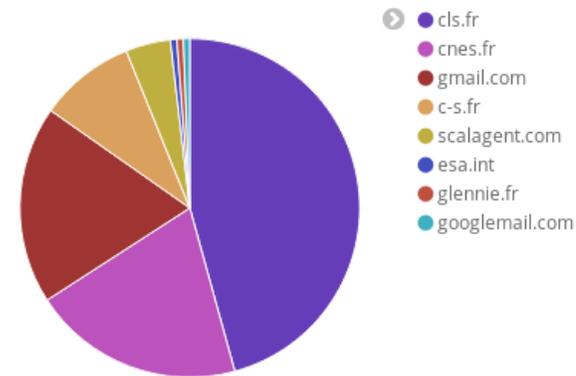
Avatar	Name	Profile	Issues	PRs
	EkicierNico	EkicierNico	2	0
	smfcrts	smfcrts	2	0
	CesarCoelho	CesarCoelho	1	0
	Jean-Christophe Malapert	J-Christophe	1	0
	LeUlukai	LeUlukai	1	0

Export: [Raw](#) [Formatted](#)

Top Authors

Name	Commits
Frédéric Briol	89

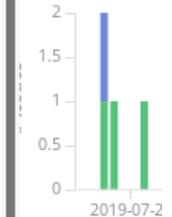
Domains



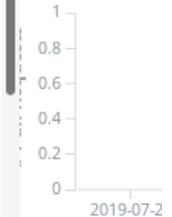
Top Repositories

Repositories	Commits	Issues	Pull Request
cnes/pangeo-pytide	15	1	1
cnes/DOI-server	12	1	0
cnes/JavaDoiApiClient	1	0	0

Issues



Pull Request



Commits



Overview

Git

Issues

Issues Timing

Pull Requests

PRs Timing

Backlog

About

■ Facilitation de partenariats et collaborations



- **Outre la valeur liée à la mutualisation qui résulte des collaborations entre acteurs, la diffusion en Open Source apporte également des gains en termes d'agilité et agit comme un facilitateur de collaborations, par la simplification des modalités juridiques et techniques de celles-ci.**
- **L'Open Source peut :**
 - Simplifier la collaboration sans être bloqué par des considérations de propriété intellectuelle
 - favorise par ailleurs les collaborations public-privé
 - initier des collaborations qui peuvent se prolonger dans d'autres contextes
 - favoriser et rendre possible une collaboration face à des partenaires susceptibles de préférer développer leurs propres solutions.

■ Valorisation financière : services associés



- Services associés au logiciel (formations, développements personnalisés, etc.)
 - Au niveau de l'établissement
 - Au niveau des autres acteurs
- **Mesure déclarative**
- Via un cadre contractuel (programme de certification)
- Vente de licences (Dual licencing)

■ Rayonnement et marque



Le choix du modèle de l'Open Source engendre généralement des conséquences réputationnelles positives, à la fois pour la **marque du projet** puisque les acteurs de la communauté participeront à la diffuser, et **pour l'établissement** en qualité d'initiateur et porteur du projet.

- Valeur intrinsèque de la marque

- L'approche par le marché
- L'approche par les coûts
- L'approche par les revenus

- Rayonnement et marque

- Mécanique d'un usage combinant la marque et la diffusion Open Source (Exclusivité de l'usage et collaboration sur le code ; Contrôle de la marque et usage de la marque)
- Modalités d'usage de la marque des projets Open Source



Freins et risques



■ Coûts associés à la mise en place d'une démarche Open Source



Le bénéfice d'une démarche Open Source induit certains coûts spécifiques associés à la mise en place :

- d'une **structure de gouvernance** et d'une **infrastructure technique** adaptées pour :
 - Produire du code collaborativement ;
 - Documenter le projet ;
 - Gérer la qualité ;
 - Interagir avec les utilisateurs.
- d'une **animation continue de** la communauté pour :
 - Stimuler la communauté
 - Prendre en compte les contributions

■ Risques associés à la diffusion en Open Source



Plusieurs risques pressentis et/ou réels doivent être anticipés :

Atteinte à la **qualité** associée à l'image de l'établissement

- **Fondement** : temps de publication courts
- **Réponse** : Mise en place des procédures internes de qualité, technique et juridique

Divulgation du « **Savoir-faire sensible** »

- **Fondement** : plus de relations et collaborations externes
- **Réponse** : sensibilisation interne et mise en place de processus spécifiques

Perte d'un levier de politique industrielle

- **Fondement** : la liberté associée au modèle (droits patrimoniaux partagés)
- **Réponse** : associer une dimension stratégique aux choix de licences à la gestion des autres actifs

Perte financière

- **Fondement** : la gratuité associée au modèle
- **Réponse** : définition de stratégie de valorisation adaptée et compensation par d'autres valorisations



Grille de critères appliquée



Grilles de critères de valorisation appliquée



	#	Éléments de valeur	Métriques	Moyens de mesure	Application à OTB	Application à Celestlab
Valeurs corrélationnelles	1	Innovation et agilité	non mesurable	S.O.	S.O.	S.O.
	2	Transparence	non mesurable	S.O.	S.O.	S.O.
	3	Qualité	- Qualité du code - Qualité juridique	Comparaison locale non praticable	S.O.	S.O.
	4	Mutualisation (Interne)	- nombre de réutilisateurs internes - nombre de contributeurs internes (nb de départements qui contribuent)		S.O.	S.O.
Développement économique	5	Nombre d'entreprises utilisant le logiciel	- Nombre total d'utilisateurs - Proportion d'entreprises	- Téléchargements - Télémétrie - Analyse des canaux de communication - Enquête d'utilisation	Significatif	Présent
	6	Répartition géographique des usages	Pourcentage d'entreprises par zones géographiques pertinentes	- Téléchargements - Analyse des canaux de communication - Enquête d'utilisation	En accord avec les objectifs du CNES	Information non disponible
	7	Services associés en aval	- Nombre d'acteurs proposant les services associés - Répartition par typologie des acteurs (PME / Grands Groupes / Établissements Publics) - répartition géographique des acteurs - CA total généré en aval	Enquête Éléments transmis dans le cadre d'une convention partenariale	Significatif	Information non disponible
Soutien à la recherche	8	Nombre d'établissements de recherche et d'enseignement supérieur utilisant le logiciel	- nombre d'établissements de recherche / Enseignement supérieur utilisateurs	- Téléchargements - Télémétrie - Analyse des canaux de communication - Enquête d'utilisation	Significatif	Présent
	9	Répartition géographique des usages	- Pourcentage d'établissements par zones géographiques pertinentes	- Téléchargements - Analyse des canaux de communication - Enquête d'utilisation	En accord avec les objectifs du CNES	Information non disponible
	10	Diffusion du logiciel dans les milieux de la recherche et de l'enseignement supérieur	- Nombre d'articles scientifiques consacrés au logiciel	Bibliométrie classique	Significatif	Information non disponible
	11	Utilisation du logiciel pour produire la recherche	- Nombre d'articles citant le logiciel pour produire la recherche	Bibliométrie adaptée à la citation de logiciel	Significatif	Présent
Mutualisation externe	12	Contributions externes	- Volume des contributions externes	Analyse statistique des outils techniques pour les différents types de contribution Recensements pour les autres domaines (communication, intervention lors de conférences, etc.)	Significatif	Non significatif
	13	Répartition des contributions	- Typologie des contributeurs (et volume associé) - Poney Factor / Elephant factor (nombre minimal d'individus/d'entités dont les contributions des employés constituent la moitié de la base de code du projet)	Filtres dans les éléments précédents Enquête	Elephant factor 1 sur le coeur / 2 sur les modules externes.	S.O.
	14	Mutualisation entre établissements publics	- part des contributions venant d'établissements publics français	Filter dans les éléments précédents	Significatif	S.O.
	15	Services gratuits pour projets Open Source	- Valeur des services utilisés	Consultation des tarifs éditeurs	Coverity fees	Non significatif
	16	Contribution du CNES a des projets tiers	- nombre de projets auquel le CNES contribue - Volume des contributions	Recensement interne	Significatif	S.O.
Monétisation des services associés	17	Valeur générée par les services fournis par le CNES en relation avec le logiciel	- CA généré	Recensement interne	Aucune	Aucune
	18	Vente de licences sur le logiciel	- nombre de licences - CA généré	Recensement interne	Aucune	Aucune
Facilitation partenariats et collaborations	19	Nombre de partenariats conclus sur la base du logiciel	- nombre de projets réalisés en partenariat entre le CNES et un autre établissement public - nombre de projets réalisés en partenariat entre le CNES et un autre établissement privé	Recensement interne	Information non disponible	Information non disponible
Rayonnement	20	Valeur de la marque en tant que telle	- notoriété de la marque	Enquête spécifique	Information non disponible	Information non disponible
	21	Utilisation de la marque par des tiers	- nombre de personnes qui utilise la marque dans la commercialisation d'un produit ou service	Enquête Recensement interne	Information non disponible	Information non disponible

■ Crédits



- Page de titre : Photo by NASA on Unsplash : <https://unsplash.com/photos/yZygONrUBe8> (domaine public)