

Mémoire de fin d'études

À l'Institut de Recherche pour le Développement (IRD)
À Bondy

Valorisation des publications open access dans un institut de recherche : numérisation et chaîne de production des documents numériques scientifiques à l'IRD.

Par SYLLA Oumar

Encadrants : Laurence Goury et Pier Luigi Rossi

Master 2 Gestion Stratégique de l'Information

Université de Paris 8 Vincennes Saint-Denis

Enseignant : Bruno Menon

2019- 2020

Remerciements

Je tiens à exprimer toute ma reconnaissance à l'ensemble de l'équipe du service IST (information scientifique et technique) de l'IRD en passant par le responsable du service Monsieur Pascal Aventurier.

Je remercie particulièrement Madame Laurence Goury et Monsieur Pier Luigi Rossi pour l'encadrement, le soutien ainsi que leurs conseils et encouragements.

J'adresse mes sincères remerciements à mon professeur encadrant Bruno Menon ainsi qu'à tous les professeurs du master Humanités numériques parcours gestion stratégique de l'information de l'Université Paris 8 Vincennes Saint-Denis.

Je remercie mes parents ainsi que mes oncles Samba, Abdoulaye et Mamadou Dramé pour le soutien constant et les encouragements.

Enfin je remercie Mouhamadou Baoumou Dramé, Bocar, Yaya et Marième Sylla.

Notice

Sylla Oumar	Valorisation des publications open access dans un institut de recherche : numérisation et chaîne de production des documents numériques scientifiques à l'IRD. Master Humanités numériques parcours gestion stratégique de l'information, 2020, p 56
Résumé Ce mémoire présente une étude sur la numérisation et chaîne de production des documents numériques scientifiques. Il permet de voir l'impact de la mise en libre accès du patrimoine scientifique de l'IRD dans les communautés scientifiques du Sud.	Mots Clés Science ouverte Bibliométrie Numérisation des documents Evaluation de la recherche Institut de recherche pour le développement (IRD) Archive ouverte institutionnelle Information scientifique et technique Analyse des besoins
Abstract This dissertation presents a study on the digitization and production chain of scientific digital documents. It shows the impact of the free access of the IRD's scientific heritage in the scientific communities of the South.	Keywords Open science Bibliometrics Scanning documents Research evaluation Development Research Institute (IRD) Institutional Open Archive Scientific and technical information Needs analysis

Plan

Introduction	5
1 La science ouverte	7
1.1 Définition et historique des mouvements Open Access.....	7
1.2 Les différents types d'open access.....	10
1.2.1 La voie verte (green road).....	10
1.2.2 La voie dorée (gold open access).....	11
1.2.2.1 Le modèle auteur-payeur.....	12
1.2.2.2 La Subvention	12
1.2.2.3 Le Freemium.....	13
1.2.2.4 La souscription.....	13
1.2.2.5 Le financement participatif (crowdfunding)	13
1.3 Le modèle hybride.....	13
1.4 Impact de l'open Access sur les publications.....	14
1.5 Quelques typologies favorables à un impact positif de l'open access.....	16
2 Les indicateurs bibliométriques	18
2.1 Les indicateurs de revues.....	19
2.1.1 Le facteur d'impact (FI)	19
2.1.2 Les quartiles	19
2.1.3 L'indice d'immédiateté (Immediacy Index).....	20
2.1.4 L'âge médian des articles cités (Cited half-life).....	20
2.1.5 L'Eigenfactor (EF)	20
2.1.6 L'Article Influence Score (AI)	20
2.1.7 Le Source Normalized Impact per Paper (SNIP).....	21
2.1.8 SCImago Journal Rank (SJR)	21
2.2 Les autres types d'indicateurs	21
2.2.1 L'AERES (Agence d'évaluation de la recherche et de l'enseignement supérieur)	22
2.2.2 Le nombre de citation	22
2.2.3 Des indices pour des institutions ou des pays	23
2.2.4 L'indice h	23
2.2.5 L'indice G	24
2.2.6 L'indice N.....	24
2.2.7 L'analyse des réseaux	24

2.2.8	Les Altmetrics	25
2.3	Mesure du taux d'accès ouvert des publications scientifiques en France	26
3	Open Access et numérisation à l'IRD	28
3.1	Le centre de Bondy	29
3.2	L'information scientifique à l'IRD et le service de numérisation	30
3.3	Les aspects techniques de la numérisation de documents	33
3.3.1	La préparation des documents	33
3.3.2	La résolution	33
3.3.3	Le mode	34
3.3.4	Les formats	34
3.3.5	Les outils utilisés dans le processus de numérisation à L'IRD	35
3.3.6	La vérification et finalisation des documents électroniques	40
3.3.7	Le reconditionnement du document	40
4	Mesurer l'impact de la mise en ligne et en accès ouvert des publications de l'IRD	41
4.1	Les statistiques de consultations	44
4.2	Exemples d'indicateurs de comparaison produits par le service de numérisation de l'IRD	48
4.3	L'enquête qualitative	51
	Conclusion	52
	Bibliographie	54
	Annexes	56
	Annexe 1 : Organigramme de l'IST	56
	Annexes 2 : Dessins open access vert et gold	56
	Annexe 3 : Enquête du service IST sur la base Horizon auprès des chercheurs de l'IRD	56

Introduction

La diffusion de l'information scientifique occupe une place très importante dans le monde. Elle présente des enjeux scientifiques, politiques et économiques. Elle est principalement assurée par un petit nombre d'acteurs : les éditeurs scientifiques, et représente une activité très lucrative. Des éditeurs comme Elsevier¹ ou Springer², Wiley Blackwell's³ et Taylor and Francis⁴ génèrent chaque année plus de 7 600 000 000 euros de chiffres d'affaires.

Les articles écrits par les scientifiques sont relus par les pairs et les articles les mieux évalués sont publiés dans les revues. Ces dernières sont vendues via des abonnements à des plateformes aux bibliothèques universitaires ou d'instituts de recherche, qui emploient ces mêmes chercheurs ayant produits les recherches scientifiques. Chaque mois 12 000 000 de chercheurs utilisent Science direct⁵, portail des revues de l'éditeur Elsevier, qui diffuse sur abonnement plus de 13 000 000 documents référencés, principalement des articles.

Aux USA depuis 20 ans le prix des abonnements aux revues scientifiques a augmenté de 7% chaque année. En 2014, les universités en France et les laboratoires de recherche ont dépensé 105 000 000 euros pour les abonnements. En Angleterre entre 2010 et 2014 les paiements ont augmenté de 50%. En Grèce en 2015, le portail public en ligne suspend ses publications par faute de moyens. Au Canada l'université de Montréal renonce à 2116 revues de Springer nature et n'en garde que 150. Ces suspensions ou désabonnements réduisent le facteur d'impact des publications issues de ces pays. Le facteur d'impact est la note attribuée à chaque journal pour évaluer sa notoriété. En fait plus un article d'une revue est cité dans d'autres recherches plus sa note augmente.

Cependant pour palier à ces contraintes financières imposées par les éditeurs privés, des luttes ont été menées pour libérer la diffusion de la recherche des circuits commerciaux. La libre circulation du savoir est un besoin naturel. La production scientifique permet d'avoir des solutions aux problèmes dans la société. Maintenant en dehors des chercheurs, les citoyens s'intéressent également à la science. C'est le mouvement pour l'accès ouvert à la publication scientifique, dont il sera question dans ce mémoire.

¹ <https://www.elsevier.com/fr-fr>

² <https://www.springernature.com/>

³ <https://www.wiley.com/en-fr>

⁴ <https://taylorandfrancis.com/>

⁵ <https://www.sciencedirect.com/>

C'est dans ce mouvement que s'inscrit l'IRD qui lutte contre les difficultés d'accès à la littérature scientifique des pays du Sud. Les universités de ces pays sont souvent confrontées aux coûts d'abonnements faramineux imposés par les éditeurs dont il est question plus haut, en situation de monopole.

Les chercheurs de l'institut de recherche pour le développement travaillent principalement avec les pays du Sud, et l'institut s'inscrit dans une dynamique d'open access pour permettre la restitution des résultats de sa recherche aux communautés concernées.

C'est dans ce cadre que son service IST, Information Scientifique et Technique a adopté une stratégie de numérisation et de mise en libre accès sur internet de la production scientifique de ses chercheurs. Dans le cadre de ce mémoire, on se posera la question de la mesure de l'impact de la mise en accès libre du patrimoine scientifique de l'IRD sur l'accès aux ressources documentaires pour les communautés scientifiques du Sud.

Après avoir présenté la science ouverte et les enjeux de l'accès ouvert aux publications scientifiques, on s'intéressera aux différents indicateurs bibliométriques qui permettent, en science, de mesurer l'impact des publications scientifiques.

On présentera ensuite le service IST de l'IRD et les opérations de numérisation qui ont été réalisées pendant le stage. Nous essaierons ensuite de répondre à la question de l'impact de la mise en libre accès des publications scientifiques sur l'accès aux ressources documentaires pour les communautés scientifiques du Sud, en nous appuyant sur les indicateurs de consultation des pdf produits par le service IST.

1 La science ouverte

1.1 Définition et historique des mouvements Open Access

Dans son ouvrage, Suber définit l'open access comme : « Né en 1991 autour du projet ArXiv, formalisé en 2001 par la Déclaration de Budapest, inscrit dans les recommandations européennes en 2012, et dans la loi française en 2016, l'accès ouvert est le mouvement de diffusion gratuite de la production scientifique sur le Web » (Suber, 2016, p. 9)

La figure ci-dessous montre les différents éléments qui tournent autour de la science ouverte à savoir : l'open process autrement dit la méthode ouverte, open data ou données de la recherche ouvertes et l'accès ouvert pour les publications.

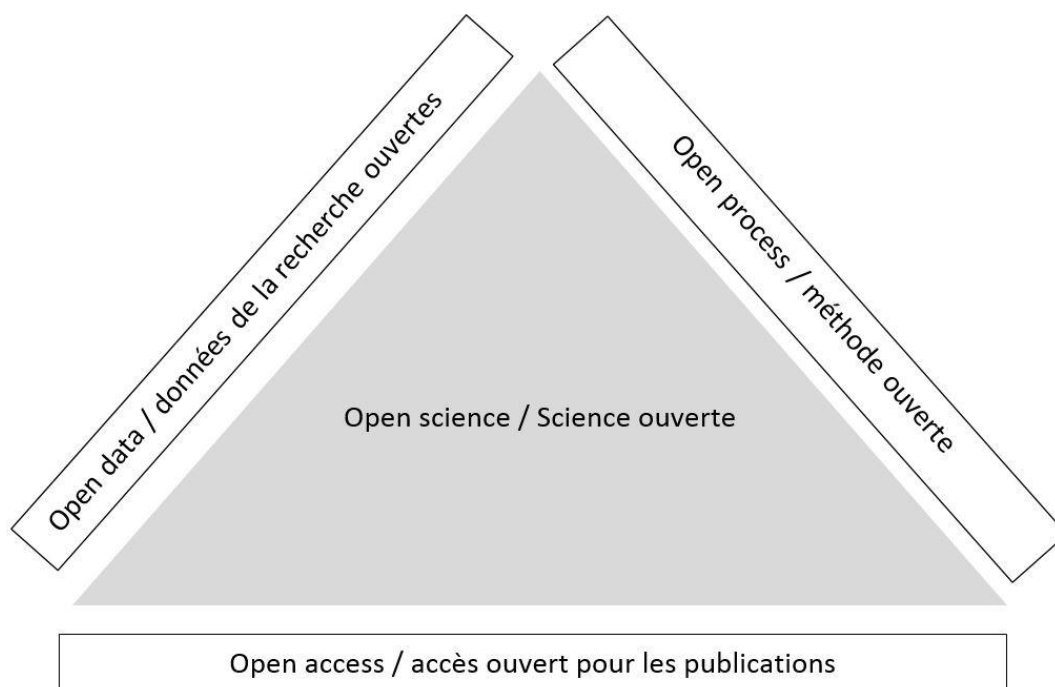


Figure 1 : Le contexte de la science ouverte et de l'open access

Source: <https://bibliotheque.u-pec.fr/science-ouverte>

L'open access permet de s'affranchir des contraintes d'accès aux documents tout en maintenant les protections du droit d'auteur sur les textes. Les textes ne peuvent pas être reproduits ou modifiés.

L'open access en outre permet au lecteur d'avoir des libertés. Parmi elles il y a le droit de diffusion publique, de partage. Cela donne plus d'ouverture aux documents en ouvrant la science aux citoyens.

Beaucoup de luttes se sont faites contre la confiscation de l'accès à la publication scientifique par les éditeurs privés. Ces luttes s'articulent autour de trois grandes déclarations (Suber, 2016, p. 7) :

- L'Initiative de Budapest⁶ acté le 14 février 2002, dont il est le principal rédacteur
- La Déclaration de Bethesda⁷ signée le 20 juin 2003
- La Déclaration de Berlin⁸ signée le 22 octobre 2003.

C'est d'ailleurs dans ce sens qu'il utilise l'expression «**BBB**» (Budapest, Bethesda, Berlin) pour parler de l'accès ouvert. On note bien que l'accès ouvert a suivi un processus qui est marqué par des dates phares qui ont permis son installation.

Pour ce qui concerne l'initiative de Budapest il faut retenir que l'accès ouvert est défini comme une libre mise à disposition sur internet de la littérature scientifique. Cette mise en liberté devra permettre à tout usager d'avoir accès au document pour le lire, le télécharger, copier et le diffuser ou l'imprimer. Cette déclaration demande que les auteurs contrôlent leurs droits d'auteurs et qu'ils soient mentionnés et cités à chaque fois.

Pour les déclarations de Bethesda et de Berlin, la définition donnée est la suivante : « Pour qu'une œuvre soit en accès ouvert, le détenteur des droits doit consentir à l'avance à ce que les usagers copient, utilisent, diffusent, transmettent et affichent l'œuvre publiquement, qu'ils créent et diffusent des œuvres dérivatives dans tout médium numérique ayant un objectif responsable, avec attribution adéquate de l'œuvre à son auteur » (Suber, 2016, p. 27).

⁶ <http://openaccess.inist.fr/?Initiative-de-Budapest-pour-l>

⁷ <http://openaccess.inist.fr/?Declaration-de-Bethesda-pour-l>

⁸ <http://openaccess.inist.fr/?Declaration-de-Berlin-sur-le-Libre>

En outre le libre accès ou open access peut être défini comme la mise à disposition gratuite, immédiate et continue sur Internet des travaux ou publications scientifiques issues de la recherche et de l'enseignement.

La science ouverte est une nouvelle approche transversale de l'accès au travail scientifique, des données et du partage des résultats de la science. Cette pratique est nouvelle et permet à la science d'être mieux valorisée et accessible à tous.

C'est d'ailleurs dans ce sens que la DIST (direction de l'information scientifique et technique) du CNRS (centre national de recherche scientifique) dira dans son livre blanc consacré à la science ouverte que : « L'open science, c'est donc un changement d'optique, qui veut se comparer à d'autres grands stades antérieurs, comme l'avènement du télescope ou du microscope. L'open science veut ainsi rendre compte des mutations qu'apportent les grandes plateformes « ouvertes » internationales : elles donnent accès à de nouveaux agencements d'acteurs de la recherche (innovation numérique, société civile), facilitées par de nouvelles démarches de partage (lois numériques) ... » (DIST, 2017, p. 12).

Les instances européennes et internationales sont favorables à une avancée de l'open science dans le monde. Beaucoup de pays mettent en place des textes qui vont dans ce sens et qui s'inscrivent dans un mouvement d'ouverture international.

C'est dans ce sens que Carlos Moedas commissaire européen pour la recherche, la science et l'innovation définit trois priorités à son action en juin 2015 à savoir : Open innovation, Open science, Openness to the world (Suber, 2016).

En France une stratégie nationale d'open science est mise en place afin de mieux partager les connaissances. Cette stratégie est portée par le conseil scientifique du CNRS⁹ (centre national de recherche scientifique).

Le gouvernement a aussi pris des positions favorables à travers un projet de loi pour une république numérique. C'est un projet qui date du 18 juin 2015. Ainsi l'open science permettra de mieux valoriser la science qui est considérée comme une priorité par le gouvernement.

Beaucoup de points de vue sont sortis entre les chercheurs et les institutions :

« La science est un bien commun à l'humanité ;

Les intérêts légitimes de préservations des secrets et de valorisation doivent être préservés ;

⁹ <http://www.cnrs.fr/>

Le texte and data mining est un droit d'observation numérique naturel nécessaire aux chercheurs dans leur démarche scientifique ;

Les clauses de cession exclusive de droit d'auteur prévues dans les contrats d'édition doivent être déclarées nulles et non avenues ». (DIST, 2017, p. 12).

1.2 Les différents types d'open access

Il existe plusieurs façons de diffuser des données de recherche ou des publications en accès libre. Elles s'appuient sur des points établis lors de grande déclaration comme celle de Berlin.

Quelques points sur la déclaration de Berlin sont :

- ✓ La diffusion complète des connaissances accessible à la société.
- ✓ L'accès libre est une source complète de connaissances humaines et de patrimoine culturel qui a été approuvée par la communauté scientifique.
- ✓ Soutenir la transition vers le paradigme de l'accès libre électronique
- ✓ Développer des moyens et des moyens d'évaluer les contributions en libre accès et les revues en ligne afin de maintenir les normes d'assurance de la qualité et de bonnes pratiques scientifiques.
- ✓ Préconiser que la publication en libre accès soit reconnue dans l'évaluation de la promotion et de la permanence.

1.2.1 La voie verte (green road)

La voie verte ou « green road » permet de faire de l'auto-archivage. Selon le contrat d'éditeur de la revue, l'auteur ou le chercheur dépose son document (pre-print, post-print) dans une archive ouverte. On appelle préprint dans le domaine de la publication scientifique, une version d'un article scientifique qui précède son acceptation par le comité de rédaction d'une revue

scientifique. Contrairement, un postprint est obtenu après un examen et acceptation par les pairs d'un article de revue de recherche sous la forme de brouillon numérique.

Une archive ouverte est une base ou plateforme où sont déposées des publications dont l'accès est libre et gratuit sur internet. C'est une plateforme en ligne qui permet le dépôt et la diffusion de contenus scientifiques avant ou après la publication dans une revue scientifique. Certaines universités mettent par exemple en place une archive institutionnelle connecté à HAL¹⁰. Ainsi les chercheurs ont la possibilité de déposer dans une archive ouverte institutionnelle comme celle de l'IRD. Ces publications sont visibles sur les deux plateformes. Cette voie permet de rendre la diffusion plus large, un archivage des contenus malgré la situation économique de l'éditeur ou de la politique. Une période d'embargo peut exister avant l'accès au texte intégral. Cette période est définie par un plan de la science ouverte de l'Union Européenne de 6 mois en sciences, techniques et médecine et à 12 mois en sciences humaines et sociales. Comme exemple on peut donner Lille Open Archive de l'Université de Lille¹¹.

1.2.2 La voie dorée (gold open access)

La voie dorée est une voie se rapprochant davantage du modèle traditionnel de publication scientifique. Ici l'article de l'auteur déjà soumis à une revue en libre accès suit un circuit classique de l'édition scientifique : La validation par les pairs des contenus (peer-reviewing), mise en page des articles à partir d'un travail éditorial ... Ainsi comme différence avec l'édition scientifique traditionnelle : le lecteur a la possibilité de consulter gratuitement les documents, sans abonnement.

En fait la publication dans une revue en open access demande souvent une prise en charge des frais de publication (APC, Article Processing Charge) par l'institution qui l'emploie l'auteur. Ici on parle de « modèle auteur-payeur ».

¹⁰ <https://hal-univ-paris8.archives-ouvertes.fr/>

¹¹ <https://lilloa.univ-lille.fr/>

Dans la voie dorée plusieurs possibilités peuvent s'offrir à un éditeur ou à une revue pour la diffusion en accès libre. Les modèles économiques actuels permettent de répondre aux coûts d'édition. Ces modèles économiques de la voie dorée ont plusieurs formes :

1.2.2.1 Le modèle auteur-payeur

Dans ce modèle l'accès au document est libre pour le lecteur. Mais l'auteur ou son institution ont des frais appelés Article processing charges (APC) pour financer la publication. Les APC ont la forme d'un coût à l'article et varient de quelques centaines d'euros à plusieurs milliers (ex. : Plos¹²) Des adhésions institutionnelles sont proposées par des éditeurs afin de bénéficier de remises sur le coût des APC. Il y a aussi des adhésions individuelles à vie qui donnent la possibilité de publier un ou plusieurs articles par an pour remplacer les APC.

Cependant il faudra être vigilant sur des points comme :

- ✓ Les éditeurs qui cherchent malgré le financement de la publication à faire profit avec des dérives tarifaires.
- ✓ Les éditeurs qui n'offrent aucune expertise (processus de relecture défaillant voire inexistant absence de comité scientifique...) d'où le nom d'éditeurs « prédateurs ».

1.2.2.2 La Subvention

Dans cette forme l'accès au document est libre. Ici le financement de la publication est pris en charge par une association, une société, un organisme, une fondation ou une institution comme le Cnrs (revue et/ou plate-forme). C'est le modèle le plus développé. Il concerne 66% des revues référencées par le Directory of Open Access Journals (DOAJ)¹³

¹² <https://plos.org/>

¹³ <https://doaj.org/>

1.2.2.3 Le Freemium

Sous cette forme l'accès reste libre pour ce qui concerne la consultation document. Cependant des services complémentaires sont payants (les fichiers PDF ou ePub, les statistiques d'usages). Comme exemple on peut donner l'exemple de OpenEdition¹⁴.

1.2.2.4 La souscription

La diffusion des ouvrages et revues se fait en fonction du seuil financier qui est visé. Il s'agit de voir si le seuil financier est atteint ou pas grâce aux contributions financières des établissements souscripteurs. A titre d'exemple on peut donner le Knowledge Unlatched¹⁵ pour les monographies et l'Open Library of humanities¹⁶ pour les revues.

1.2.2.5 Le financement participatif (crowdfunding)

Le modèle repose sur le fait que les projets potentiels soient mis en ligne. Un montant est fixé et une durée afin de financer le projet. Cela permet à une communauté en général de choisir et de financer un projet par des dons financiers. Avec suffisamment de soutien financier, un projet peut couvrir ses coûts de production et être publié.

1.3 Le modèle hybride

Dans le modèle hybride l'abonnement est payant. Ici l'auteur peut publier en libre accès son article en payant des frais. Ces revues sont souvent éditées par de grands éditeurs commerciaux. C'est un modèle remis beaucoup en cause parce que l'institution paye plusieurs fois pour l'accès au document (abonnement + frais pour la publication en open access si ses auteurs en font le choix). Cela crée un débat important dans la communauté scientifique.

¹⁴ <https://www.openedition.org/>

¹⁵ <https://www.knowledgeunlatched.org/>

¹⁶ <https://www.openlibhums.org/>

1.4 Impact de l'open Access sur les publications

La notoriété d'un article d'une revue augmente plus s'il est cité dans d'autres recherches. Cela augmente la note de l'article. Donc la citation est ici considérée comme l'indice de qualité d'un article scientifique : plus un article est cité plus il est considéré comme étant de bonne qualité.

Beaucoup d'études ont été publiées pour mesurer l'impact de l'open access sur la quantité de citation des publications. Un article en open access est plus accessible et aura donc plus de chance d'être cité.

Ainsi beaucoup de débats tournent autour de cette question. Certaines études montrent une grande corrélation entre les citations et le nombre de téléchargements et d'autres font état de facteurs liées aux caractéristiques des articles.

C'est dans ce sens que dans leur article, Bordignon et Andro (Bordignon & Andro, 2016) montrent à partir de leurs études sur les revues open access que les articles avec des titres courts sont plus vus et ont plus de citations contrairement à ceux avec des titres longs.

D'autres critères interviennent également dans le taux de citation d'un article : les articles comportant des titres avec des points d'interrogation, une référence à une région géographique particulière, deux-points ou un tiret reçoivent un faible nombre de citation.

En outre ils ajoutent le fait que les articles les articles décrivant une méthode sont moins cités que ceux qui ont un titre qui décrivent des résultats. Les financements sponsorisés permettent aussi à des publications d'être plus citées.

Par ailleurs d'autres critères comme le domaine disciplinaire, le statut des auteurs, le genre, le prestige de l'institution, le pays de résidence, la longueur des papiers (citation de beaucoup de références et les reviews), la revue, le nombre d'équations sur le texte peuvent influencer sur les citations reçues.

Les résultats, dans l'article de Bordignon et Andro (Bordignon & Andro, 2016), sont comparés en fonction de différents éléments. De ce fait on peut avoir une comparaison portant sur :

- ✓ Un domaine disciplinaire ou plusieurs domaines

- ✓ Les documents dans une archive disciplinaire ou institutionnelle, Google Scholar, Scopus, Web Of science, arXiv etc
- ✓ Une revue entre articles non open access et articles open access avec le nombre de citations
- ✓ Le type de documents (ouvrages, chapitres d'ouvrages, les versions d'une publication, working papers)
- ✓ La notoriété des revues (facteurs d'impact, etc.)
- ✓ Les revues d'une même langue ou d'un même éditeur
- ✓ Une région géographique
- ✓ La répartition sur des périodes différentes sur des volumes différents

Bordignon et Andro (Bordignon & Andro, 2016) parlent du projet Opcit (The Open Citation project)¹⁷ qui a permis de mettre en place de 2004 à 2013 une bibliographie sur les publications scientifiques en montrant l'impact de l'open access sur le nombre de citations.

Avec des résultats réalisés sur 70 études, 46 sont favorables à l'open access et montrent son avantage. Tandis que 17 sont défavorables et ne voient pas l'avantage de l'open access sur les citations. Le reste des études à savoir 7 sont non tranchées sur la question de l'open access. Un article publié en 2001 (Lawrence, 2001) mettait en évidence à partir d'un échantillonnage de 119 924 articles de conférences dans le domaine des sciences informatiques un accroissement de 157% des citations de publications pour les articles en open access. D'autres études multiples dans d'autres domaines arrivent à des conclusions identiques.

Ce qu'il faut noter c'est que la plupart des études montrent que l'open access a un effet positif sur le taux de citations des articles.

¹⁷ <http://opcit.eprints.org/>

1.5 Quelques typologies favorables à un impact positif de l'open access

Pour un impact positif de l'open access sur les documents, Bordignon et Andro (Bordignon & Andro, 2016) utilisent un certain nombre de facteurs que nous présenterons ci-dessous.

Early Advantage ou avantage de précocité

Le fait de rendre publique plus tôt une recherche permet logiquement son utilisation plus tôt. L'impact de la recherche augmente, à partir de l'auto-archivage de preprints avec une possibilité d'être étudié ou cité, avant son évaluation et sa publication. C'est le cas de travaux déposés sur arXiv¹⁸ avec une première forme. Cela peut accélérer la citation grâce au fait que les papiers soient accessibles plus tôt sur arXiv.

Quality Advantage ou avantage de qualité

L'augmentation de l'impact sur la recherche passe aussi par l'auto-archivage de postprints après acceptation de la publication. Cela est due au fait que cette version, revue par les pairs, est de meilleure qualité et accessible librement. Donc la qualité joue un rôle très important sur le nombre de citations.

Quality Bias ou avantage lié à l'auto-sélection

Le fait que les auteurs eux-mêmes choisissent leurs meilleurs papiers pour l'auto-archivage joue aussi un rôle important sur l'impact ou la différence dans le taux de citation.

Usage Advantage ou avantage d'usage ou de téléchargements

Les téléchargements augmentent grâce à l'auto-archivage. Et de ce fait cela augmente l'impact sur la recherche. Les téléchargements permettent de voir et de se faire une idée sur le nombre de citations pour une recherche.

¹⁸ <https://arxiv.org/>

Competitive advantage ou avantage concurrentiel

Le fait que les articles en open access sont plus accessibles que ceux qui ne le sont pas. De ce fait ils sont plus cités et gagnent plus en visibilité.

En résumé on notera bien dans ce chapitre que l'open Access joue un rôle très important et donne un effet positif sur le taux de citations des articles. La science ouverte donne également plus de libertés aux lecteurs en rompant avec les contraintes financières imposées par les éditeurs privés et d'accès aux documents.

2 Les indicateurs bibliométriques

La bibliométrie permet de mesurer et d'évaluer, la production et la diffusion d'articles ou autres types de publications. Elle se base sur une méthode statistique et mathématiques. Elle est une partie des sciences de l'information grâce à la bibliométrie, on peut évaluer la recherche sur plusieurs niveaux : pays, continent, laboratoire, établissements, chercheur, équipe etc. Nous nous intéresserons aux indicateurs bibliométriques dans la mesure où l'IRD produit des indicateurs alternatifs (indicateurs de consultation des PDF) pour mesurer l'impact de la mise en ligne de ses publications.

Une analyse portant sur la littérature scientifique publiée entre 1850 et 1860 sur l'anatomie est considérée comme étant la première étude bibliométrique (Cole & Eales, 1917).

Afin de mesurer l'impact d'un article on s'appuie sur son usage, l'évaluation par les pairs, le taux de citation dans les outils bibliométriques et les altmetrics.



Figure 2 : Différentes façons de mesurer l'impact d'un article

Source : <http://altmetrics.org/manifesto/>

Le calcul des indicateurs s'appuie sur les bases bibliographiques dans lesquelles les articles sont signalés. Elles doivent être mentionnées. Sur le résultat de ce calcul d'autres éléments influent sur le résultat comme : l'antériorité des citations, corrections apportées, Les sources retenues dans la base et les types de publication.

2.1 Les indicateurs de revues

Les indicateurs permettent de voir si une revue est bien notée ou pas. Ils montrent surtout à partir des citations, si les articles d'une revue sont cités ou pas.

2.1.1 Le facteur d'impact (FI)

Il mesure pour une revue la fréquence de citations pour les articles. Il permet de voir si les articles sont cités dans d'autres revues sur une année donnée. Cet indicateur concerne plus des revues en sciences de l'ingénieur ou sciences dures.

Il se calcule pour l'année N en divisant le nombre de citations (pendant l'année N) de ses articles publiés les années N-1 et N-2 par le nombre d'articles publiés dans la revue les années N-1 et N-2.

Il est mis en parallèle avec le facteur d'impact à 5 ans qui se calcule de la même façon en suivant les cinq années précédentes. Ce facteur permet de juger la croissance, la stabilité ou la diminution pour la revue.

Pour toute revue, le facteur d'impact se compare avec d'autres facteurs d'impact des revues dans le même domaine. Sa publication se fait dans le Journal Citation Report (JCR) ¹⁹.

2.1.2 Les quartiles

C'est un indicateur créé par l'Institut National de Recherche Agronome (INRA)²⁰. Certaines revues sont présentes dans plusieurs domaines, et chaque facteur d'impact est comparé avec les facteurs d'impact des revues du même domaine.

¹⁹ <https://jcr.clarivate.com>

²⁰ <https://www.inrae.fr>

Le classement peut se faire dans le quart des revues qui ont un facteur d'impact le plus élevé (Q1), ou être dans le quart inférieur (Q2) ou dans la seconde moitié des revues (Q3 et Q4).

2.1.3 L'indice d'immédiateté (Immediacy Index)

Cet indice calcule le rapport entre le nombre de citations obtenues par la revue l'année N sur le nombre d'articles publiés par cette revue l'année N. Il permet de savoir si les articles d'une revue sont vite cités ou non.

2.1.4 L'âge médian des articles cités (Cited half-life)

Il permet de savoir le nombre d'années au bout duquel le nombre cumulé de citations pour les articles de la revue atteignent 50% du nombre total de citations pour la revue.

Cet indice permet à une bibliothèque de mieux gérer sa politique d'archivages pour les revues papiers.

2.1.5 L'Eigenfactor (EF)

Cet indice est basé sur le nombre de citations des articles d'une revue sur 5 ans. Il se centre sur le même corpus de revues que le facteur d'impact. Leur différence vient du fait qu'il ne prend pas en compte les auto-citations et ajoute une pondération aux citations parues dans la revue qui sont fortement citées.

2.1.6 L'Article Influence Score (AI)

Il calcule l'influence moyenne des articles d'un journal. Ce calcul se fait sur une période de 5 ans. Le calcul se fait en divisant l'Eigenfactor par la part des articles parus dans cette revue en le rapportant à l'ensemble des articles dans toutes les revues du corpus. L'influence moyenne

est de 1. Un journal qui a une note supérieure à 1, signifie que ses articles ont une influence supérieure à la moyenne des revues de la catégorie.

2.1.7 Le Source Normalized Impact per Paper (SNIP)

Son calcul se fait en divisant le nombre de citations obtenues par les articles d'une revue R durant les 3 années précédentes par le nombre d'articles publiés par la revue pendant les 3 ans, et le potentiel de citations de cette revue dans son domaine.

On peut définir le potentiel de citation comme la fréquence à laquelle les articles du domaine citent d'autres articles publiés dans les 3 années précédentes. Une revue médiane normale a un potentiel de citation de 1.

2.1.8 SCImago Journal Rank (SJR)

C'est un indicateur qui ressemble au facteur d'impact d'une revue. Il se base sur le nombre de citations obtenues par les articles d'un journal. Dans son calcul plus complexe, l'autocitation (c'est-à-dire le fait qu'un auteur cite ses propres articles) n'est pas prise en compte et on pondère chaque citation par le prestige de la revue citante.

2.2 Les autres types d'indicateurs

Il existe d'autres types d'indicateurs qui permettent de faire des classements et le comptage des publications. Ces indicateurs sont aussi très importants dans le domaine de la publication scientifique et sont pris en compte.

2.2.1 L'AERES (Agence d'évaluation de la recherche et de l'enseignement supérieur)

L'AERES propose un classement des publications en suivant plusieurs critères. Parmi ces critères on peut lister :

ACT : Communications avec actes dans un congrès international ou national

COM : Communications orales sans actes dans un congrès international ou national

AFF : Communications par affiche dans un congrès international ou national

ACLN : Articles dans des revues avec comité de lecture non répertoriées dans des bases de données internationales.

ACL : Articles dans des revues internationales ou nationales avec comité de lecture répertoriées dans les bases de données internationales

OS : Ouvrages scientifiques (ou chapitres de ces ouvrages)

OV : Ouvrages de vulgarisation (ou chapitres de ces ouvrages)

ASCL : Articles dans des revues sans comité de lecture,

INV : Conférences données à l'invitation du Comité d'organisation dans un congrès national ou international.

DO : Directions d'ouvrages

AP : Autres publications

TH : Thèses soutenues

2.2.2 Le nombre de citation

Le comptage des citations pour un auteur donné s'appuie sur :

- Le nombre total de citations des articles : Ici on fait la somme des citations reçues par toutes les publications

- Le nombre moyen de citations par article (taux de citation) : Le calcul se fait en divisant le nombre total de citations reçues par ses publications par le nombre total de ses publications.

2.2.3 Des indices pour des institutions ou des pays

Pour évaluer la recherche dans un pays ou les établissements de recherche, des indices particuliers sont prises en compte.

Parmi ces indices nous pouvons citer :

La part des publications : ici on peut donner l'exemple de la part des publications de la France dans l'Union Européenne. Elle se calcule en pourcentage en divisant la part de publications de la France sur le nombre de publications de l'Europe, fois 100.

L'indice de Spécialisation scientifique : Il calcule la part de publication d'une entité dans un domaine par rapport à sa contribution globale. Exemple la part de publication de la France en géologie dans l'Union Européenne par rapport à sa contribution globale dans l'UE.

La part de citation à deux ans, qui mesure pour une année donnée le nombre de citations reçues sur l'année N et N+1, le nombre total des citations des établissements français à l'année N et N+1, fois 100

L'indice d'impact relatif à deux ans, qui calcule pour un établissement sur 2 années la part des citations reçues sur la part des publications de cet établissement au niveau français.

2.2.4 L'indice h

L'indice H est aussi appelé **indice de Hirsch** ou index et **h index** en anglais. Il s'intéresse pour une unité à l'impact d'un scientifique ou la productivité scientifique. Il s'appuie sur le niveau de citation de ses publications.

L'indice H de N veut dire qu'un auteur a au moins N publications cités N fois.

2.2.5 L'indice G

Il est une variante de l'indice H. L'indice G est toujours supérieur. C'est en fait le rang le plus élevé qui concernent le G premières publications cumulant un nombre de citation minimum de G^2 .

Exemple : Si un chercheur avec 18 de ses publications plus citées cumule 369 citations, alors son nombre de citations est donc d'au moins 18^2 qui est égale à 324 pour le rang G^2 .

2.2.6 L'indice N

L'indice N permet de rendre compte des années d'exercice. Il est aussi appelé indice N. Son calcul se fait en divisant l'indice H par le nombre d'années à partir de la première publication. Un chercheur qui a par exemple un indice H égale à 4 et dont son premier article date de 2 ans, aura un indice N de 2.

2.2.7 L'analyse des réseaux

L'analyse de réseaux est un facteur important pour les études bibliométriques. Elle permet de cartographier des relations entre différents éléments comme :

- ✓ Mots du titre
- ✓ Mots clés
- ✓ Auteurs
- ✓ Citations
- ✓ Mots extraits de métadonnées des publications
- ✓ Co-publications dans un domaines entre unités de recherche
- ✓ Des thématiques et des acteurs
- ✓ Informations géographiques

2.2.8 Les Altmetrics

Les indicateurs appelés ‘altmetrics’ permettent de mesurer l'influence de chaque article sur le web social.

Les altmetrics se composent :

- ✓ Du nombre de téléchargements ou de vues dans la base où l'article est déposé
- ✓ Du nombre de captures faites par les utilisateurs dans les outils de bookmarking ou de référence management²¹ .
- ✓ Du nombre de citations dans Google Scholar²² et Wikipedia²³
- ✓ Du nombre de mentions/citations sur le web social :
 - ✓ Certains réseaux sociaux comme ResearchGate²⁴
 - ✓ Les blogs hébergés par Google²⁵
 - ✓ Twitter²⁶

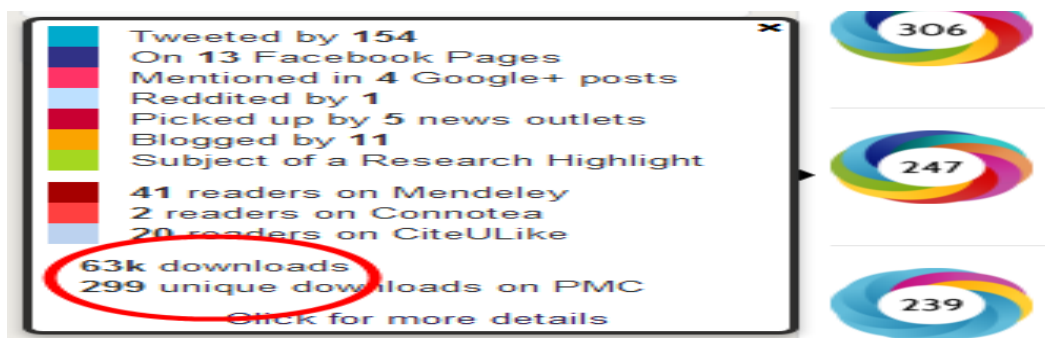


Figure 3 : Mesure de l'impact sur le web social d'un article par Altmetrics.org

Source : <http://altmetrics.org>

²¹ <https://www.mendeley.com>

²² <https://scholar.google.fr/>

²³ <https://fr.wikipedia.org>

²⁴ <https://www.researchgate.net/>

²⁵ <https://www.google.com/>

²⁶ <https://twitter.com/>

2.3 Mesure du taux d'accès ouvert des publications scientifiques en France

L'Observatoire des sciences et techniques (OST) du Hcéres²⁷ annonce dans un communiqué du 09 juillet 2020 une publication qui s'intitule : « Mesurer le taux d'accès ouvert des publications scientifiques : le cas de la France ».

L'étude porte sur une comparaison de différents calculs sur le taux d'accès ouvert des publications. Elle identifie également les sources de résultats pour la France. Elle se centre sur un indicateur normalisé qui se calcule pour différents pays et pour plusieurs types d'institutions françaises. Elle se compose de quatre grandes parties qui porte sur :

- ✓ Une analyse rapide des disparités de repérage des publications en accès ouvert (AO) selon les sources utilisées et les types d'accès ouvert.
- ✓ Une comparaison des différentes estimations du taux de publication en AO pour la France et identifie les choix méthodologiques qui influencent les résultats obtenus
- ✓ Un indicateur d'accès ouvert normalisé par spécialité disciplinaire qui permet de comparer différentes institutions ou pays sans biais dû à la composition disciplinaire de leurs publications
- ✓ Une comparaison de taux d'accès ouvert brut et l'indicateur normalisé, pour les 20 premiers pays producteurs de publications et pour les institutions françaises

Source : <https://www.ouvrirlascience.fr/mesure-du-taux-dacces-ouvert-des-publications-scientifiques/>

²⁷ Hcéres : Haut Conseil de l'évaluation de la recherche et de l'enseignement supérieur

En résumé on note que différents indicateurs sont pris en compte pour l'évaluation des publications scientifiques. Ces indicateurs bibliométriques se basent sur des calculs statistiques et mathématiques. C'est d'ailleurs dans ce sens que l'IRD produit également différents types d'indicateurs bibliométriques afin d'évaluer les publications scientifiques et son impact sur la production scientifique de ses chercheurs, mais également des indicateurs 'alternatifs' pour mesurer l'impact de la mise à disposition en open access de la production scientifique de l'institut, ce que l'on analysera dans le chapitre suivant.

Source : <https://formadoct.doctorat-bretagneoire.fr/bibliometrie>

3 Open Access et numérisation à l'IRD

Le stage s'est déroulé au sein du service IST de l'IRD. Le service IST s'occupe du recensement de toutes les publications produites par les chercheurs de l'IRD. Il dépend de la mission culture scientifique et technologique (MCST), directement rattachée au pôle science (voir organigramme en annexe). Cette mission est chargée d'assurer la diffusion et le partage des résultats de la recherche, et joue un rôle très important dans la relation entre science et société. Le service IST, par ses différentes activités, vient également en appui aux équipes de recherche. Il se donne comme tâche de fournir aux chercheurs IRD les ressources documentaires : revues scientifiques, services en ligne, bulletin de veille, bases de données bibliographiques, centres de documentation etc. Le fonds papier de l'IRD contient plus de 100 000 documents conservés sur le site de Bondy. La base de données Horizon / Pleins textes²⁸ compte plus de 83 000 documents PDF dont 67 000 sont en libre accès.

La publication scientifique renvoie à plusieurs types de documents, rédigés par les chercheurs, qui sont scientifiques ou techniques examinées de façon scientifique par un comité de lecture pour validation par les pairs. On peut citer comme types de publications :

- ✓ Les revues scientifiques
- ✓ Les comptes rendus de congrès scientifique
- ✓ Les ouvrages collectifs avec des articles de revue ou de recherche autour d'un thème donné produits par plusieurs chercheurs voire éditeurs
- ✓ Les monographies portant sur un thème de recherche

A l'IRD les unités mixtes de recherche sont appelées à produire des listes de publications. Cela peut se faire dans plusieurs cadres comme : les évaluations avec le rapport HCERES, les demande d'attribution de moyens, appels à projets, production d'indicateurs ou de bibliométrie. En 2013 Nicolas Barts (Barts & Cavet, 2014) dans son rapport sur l'année 2013 et les évolutions entre 2004 et 2013²⁹ souligne quelques faits marquants sur la publication scientifique à l'IRD :

²⁸ <https://horizon.documentation.ird.fr>

²⁹ http://horizon.documentation.ird.fr/exl-doc/pleins_textes/divers15-03/010064292.pdf

- ✓ 3682 publications ont été réalisées en 2013 par l'ensemble des unités de l'IRD (Périmètre IRD UMR)
- ✓ 1 476 publications ont été signées en 2013 par les scientifiques de l'IRD (périmètre IRD strict), soit 3 % de plus qu'en 2012.
- ✓ Soit une légère augmentation annuelle par rapport à celle de l'année précédente.
- ✓ Les scientifiques de l'IRD sont donc présents comme auteurs dans 40 % des Publications des UMR.
 - ✓ Ce taux de présence est très différent selon les regroupements thématiques : 62 % en Santé mais seulement 20 % en géosciences.
 - ✓ Au sein du périmètre IRD strict, 22 auteurs ont signé au moins 10 publications en 2013, et 134 ont signé au moins 5 publications.

3.1 Le centre de Bondy

Le centre est situé historiquement à Bondy. Il est devenu en 2017 la délégation régionale Ile de France et met en œuvre des missions et des décisions en étant plus près des 22 unités mixtes de recherche et des partenaires qui se trouvent dans la région de Paris. Elle développe beaucoup de partenariats avec des universités, mais également avec des acteurs économiques et associatifs, et des collectivités territoriales. Elle joue un rôle très important en animation des communautés scientifiques du Nord et du Sud avec des référents scientifiques.



Figure 4 : Les partenaires de L'IRD

Source : <https://www.ird.fr/>

3.2 L'information scientifique à l'IRD et le service de numérisation

L'IRD dispose de nombreux centres de documentation, présents en France et à l'étranger. Ils sont équipés de bases de données interne avec des outils documentaires informatisés. La recherche à l'IRD est orientée vers le développement et en coopération avec les pays du Sud. De ce fait les centres de documentation jouent un rôle très important en fournissant aux chercheurs des ouvrages, périodiques, articles et de la littérature grise afin d'alimenter leurs recherches. Ils permettent également de valoriser les publications des chercheurs l'IRD.

Cette préoccupation est présente à l'IRD depuis longtemps. Ce processus repose sur des articles du décret constitutif de l'institution, tiré du rapport présenté par le Professeur Raoul COMBES (*Office de la Recherche Scientifique et Technique Outre-Mer : organisation - activités : 1944-1955*, s. d.), membre de l'institut directeur de l'office de la recherche scientifique et technique Outre-Mer.

Ce rapport a été présenté aux : au Ministre de la France d'Outre-Mer, président du conseil d'administration de l'office de la recherche scientifique et technique Outre-Mer ; aux membres

du comité directeur du fonds d'investissement pour le développement économique et social des Territoires d'Outre-Mer et aux membres du conseil d'administration du fonds commun de la recherche scientifique et technique Outre-Mer.

Nous donnerons ci-dessous un aperçu des articles fondateurs de la mission du service IST de l'IRD :

ART. 6. - Le Centre de documentation est chargé d'assurer le dépouillement, la conservation et la diffusion de la documentation scientifique et technique se rapportant aux activités de l'O.R.S.T.O.M.

Raoul COMBES, office de la recherche scientifique et technique Outre-Mer, organisation - activités 1944-1955, annexe V, P.151-152

ART. 2.

C. De procéder à la constitution d'archives, de bibliothèques, de collections scientifiques et de la documentation nécessaire à l'étude des questions intéressant l'Afrique équatoriale française et l'Afrique centrale en général.

Raoul COMBES, office de la recherche scientifique et technique Outre-Mer, organisation - activités 1944-1955, Annexe IX, P.164

ART. 2.

3. De procéder dans le même domaine à la constitution d'archives, de collections et de documentation scientifiques.

Raoul COMBES, office de la recherche scientifique et technique Outre-Mer, organisation - activités 1944-1955, Annexe X, P.16

En 1990, Jean-Paul Roux-Fouillet (Lubkov & Roux-Fouillet, 1990, p. 34), alors responsable du secteur documentaire lors de son interview sur la question du rôle de la documentation à l'Orstom (ancienne appellation de l'IRD) précise que : « Il est très important, il ne faut pas oublier que l'Orstom réalise essentiellement de la recherche fondamentale. 50 % du personnel est composé de chercheurs. De plus, nous comptons un très grand nombre d'ingénieurs et de

techniciens. Leurs travaux s'exercent sur tous les continents...Pour toutes ces missions, les chercheurs ont besoin d'une importante documentation ».

L'IRD a commencé la numérisation du fonds documentaire patrimonial depuis 1996. Le service de la numérisation de l'IRD est reconnu pour sa maîtrise dans le domaine de la numérisation des documents scientifiques. Des échanges se font souvent dans le cadre de partenariat pour former du personnel de numérisation dans les pays en développement.

Par ailleurs, le service de numérisation développe des partenariats avec les institutions universitaires des pays du Sud à travers des mises en accès de productions scientifiques.

C'est d'ailleurs dans ce sens qu'en 2004 un projet de création d'une bibliothèque électronique au Burkina Faso, donnait accès à 120.000 pages de documents scientifiques réalisées par treize institutions (Rossi, 2005).

Dans son article (Rossi, 2018), l'auteur montre que la première formation avait eu lieu à l'Institut Sénégalaise de Recherche Agricole (ISRA) en 2001 à Dakar. D'autres formations ont eu lieu par la suite : « Plusieurs autres interventions ont été réalisées par la suite, notamment en Tunisie et au Burkina Faso entre 2003 et 2005. En 2008-2009 l'IRD a assuré la gestion du projet SIST (Système d'information scientifique et technique) pour le volet numérisation de documents. En 2013-2014 l'IRD a copiloté le projet « Thèses en Afrique » avec le CAMES (Burkina Faso) et l'IMIST (Maroc). Ces différents projets nous ont permis de former 194 professionnels de l'IST dans 13 pays africains et d'installer 45 ateliers de numérisation. » (Rossi, 2018b, p. 1).

En plus de la numérisation des documents, le service fait la transposition des fichiers informatisés vers le format PDF et aide à la création de bibliothèques numériques.

Lors de la mise en place du projet de numérisation en 2004 au Burkina Faso, environ 10.000 à 15.000 ont été traités. Ces travaux ont été réalisés dans deux ateliers à savoir le CNRST (Centre national de la recherche scientifique et technique) et à l'EIER/ETSHER (Ecole d'ingénieurs de l'équipement rural / Ecole des techniciens supérieurs de l'hydraulique et de l'équipement rural).

Nous présenterons ci-dessous les deux principaux outils de diffusion de la production scientifique développés par le service numérisation de l'IRD.

3.3 Les aspects techniques de la numérisation de documents

La numérisation est l'action de convertir une information issue d'un support physique en codage numérique. Elle se fait en mode binaire (en 0 et 1) sur plusieurs bits. Un bit est la plus petite unité de traitement de l'information dans un ordinateur : 1024 bits font un kilo-octet (ko) et 1024 KO égalent un méga-octet. La lecture de fichiers numériques se fait par des logiciels, périphériques etc. Il existe plusieurs phases dans le processus de numérisation des documents.

3.3.1 La préparation des documents

Cette phase consiste à préparer le document à numériser. Elle consiste d'abord à massicoter le document présentant une reliure rigide ou le désassemblage des reliures légères. Cette opération permet d'avoir des feuilles séparées pour éviter d'omettre ou de sauter des pages. Le traitement numérique se fait ensuite avec l'aide d'un scanner. Il existe différents types de scanners : A2, A3 etc.

3.3.2 La résolution

La résolution porte sur la taille des pixels. Elle se calcule en Pixels par pouce (PPP). En fait plus la résolution est élevée, plus les pixels sont petits. De ce fait il en faut plus pour remplir la matrice d'un document de même dimension.

1 pouce x 1 pouce à 12 pixels par pouce = 12 pixels x 12 pixels = 144 pixels

1 pouce x 1 pouce à 24 pixels par pouce = 24 pixels x 24 pixels = 576 pixels

A chaque fois en doublant on se retrouve avec deux fois plus de pixels en largeur et en hauteur.

3.3.3 Le mode

A chaque fois qu'on numérise un document on devra choisir un mode approprié aux caractéristiques du document à traiter. Pour une image en noir et blanc il varie de deux valeurs possibles, en nuances de gris à 256 valeurs et de 16 millions en couleur.

En noir et blanc le mode permet aux pixels d'avoir la valeur : noir ou blanc. Ce mode est très compact. Ce mode est composé d'un bit pour déterminer la valeur de chaque pixel.

On utilise le mode niveau de gris pour la numérisation de photographies en noir et blanc et pour des illustrations monochromes comprenant des nuances. Ici on a 8 bits pour la définition de la valeur de chaque pixel.

Le mode de couleurs permet d'utiliser la couleur en s'approchant de la capacité humaine de perception des couleurs. Ici on utilise 24 bits pour la numérisation des 16 millions de valeurs de couleurs.

3.3.4 Les formats

Il existe différents types de formats dans la numérisation des documents. Parmi ces formats nous avons :

Le format EPS (Encapsulated Postscript) : Il est composé de deux parties distinctes à savoir document comme tel et une version basse résolution pour l'affichage de l'image dans un logiciel de mise en page. Il est le moins compact de tous les formats sauf en utilisant la compression destructive optionnelle.

Le format TIFF (Tagged Image File Format) : Il est compatible avec les logiciels de la chaîne graphique. Il permet en option d'avoir l'application d'une compression destructive ou non destructive au document pour le rendre moins lourd.

Le format PDF (Portable Document File) : Il s'appuie sur un langage Postscript. Il est populaire chez les utilisateurs et le grand public. Il s'adapte à différents besoins et offre toute forme des fonctions de compression.

Le format JPEG (Joint Photographic Expert Group) : Ce format permet la compression de documents photographiques. C'est un format très recommandé pour le téléchargement d'une image par internet.

Le format GIF (Graphic Interchange Format) : Il est accessible aux images à la résolution de 72 PPP et en mode couleurs. Il permet de faire une compression non destructive et rend bien compte du nombre de couleurs utilisées dans l'image.

Le format PNG (Portable Network Document) : C'est un format libre en droit ; Le Format PNG 8 permet de faire la même chose que le format GIF. Et le format PNG 24 permet la conservation des documents couleurs.

A l'IRD le format utilisé pour la numérisation des documents est le format PDF. Il concerne les documents papiers et numériques de types Word ou PAO. Ces documents sont ensuite passés à l'OCR (reconnaissance optique des caractères) qui permet d'indexer de façon intégrale le contenu des documents sur les moteurs de recherche Google.

3.3.5 Les outils utilisés dans le processus de numérisation à l'IRD

Le service de numérisation procède de différentes manières dans la sélection des documents. Les documents peuvent être sélectionnés en fonction du nombre d'exemplaires ou du domaine : exemple la pédologie.

Les documents sont ensuite numérisés soit sur place à Bondy ou envoyés chez un prestataire. Ce dernier suit et doit respecter un cahier de charges qui est défini par le service de numérisation. Cela rentre dans le but d'assurer la qualité des documents numérisés en respectant le poids des PDF qui sont destinés à être téléchargé dans les pays où les performances de connexion internet sont limitées.

Les documents numérisés à Bondy sont dans le cadre d'une réponse à une demande d'un usager et de façon ponctuelle.

Pour la numérisation des documents on travaille essentiellement avec des scanner A0, A4 et A3. Ces derniers permettent de scanner à grand nombre les documents des chercheurs de l'IRD et de la littérature grise.

Le travail se fait avec les logiciels Adobe Acrobat professionnel et Adobe Photoshop CS2³⁰.

³⁰ <https://www.adobe.com/fr/>

La numérisation de document demande beaucoup de concentration et de vérification afin de réaliser à bien les tâches. Le document doit être lisible et bien numérisé en respectant le bon mode.

Quelques illustrations de numérisation avec Adobe Acrobat et Adobe Photoshop CS2 :

Pour numériser la page ci-dessous on a utilisé le mode image noir et blanc avec une résolution de 300 dpi. On s'assure que le chargeur papier est bien en réglages détection des doubles pour ne pas sauter des pages. Le seuil utilisé est le seuil statique qui est bien approprié pour la numérisation des documents en noir et blanc. Le document est numérisé en recto seul car ne présentant pas de texte en verso.

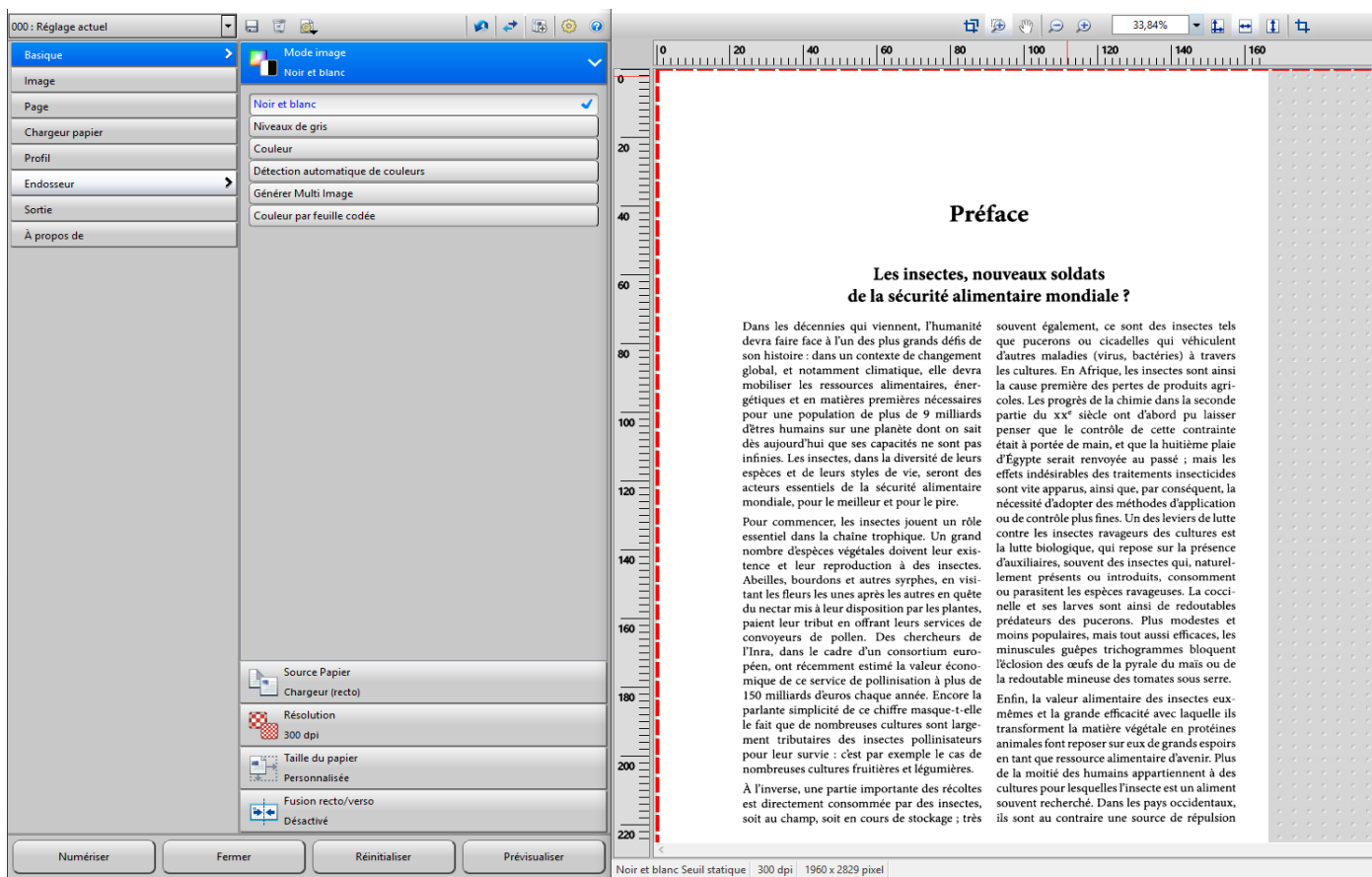


Figure 5 : Prévisualisation d'un document au bon mode avant numérisation sur Adobe Acrobat professionnel

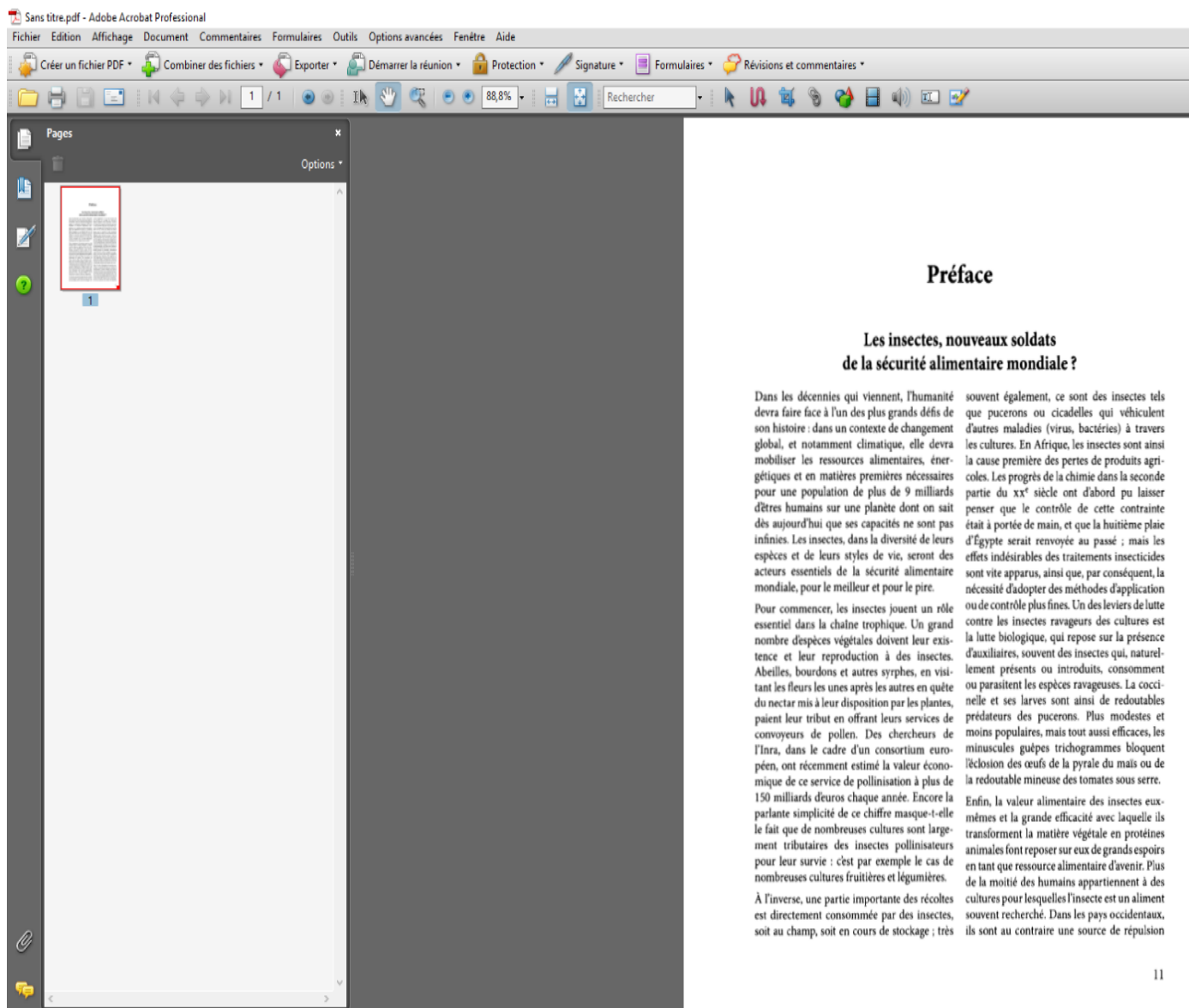


Figure 6 : Page d'un document numérisé avec Adobe Acrobat professionnel

La numérisation de cette page de couverture est faite sur Adobe Photoshop. Cette méthode permet de nettoyer la couverture et de pouvoir la rogner par la suite. On rogne la couverture pour pouvoir séparer la première et la quatrième de couverture. Le mode utilisé est la couleur à 24 bits avec une résolution de 300 dpi. Le document est numérisé avec un scanner A3 car la couverture est très rigide.

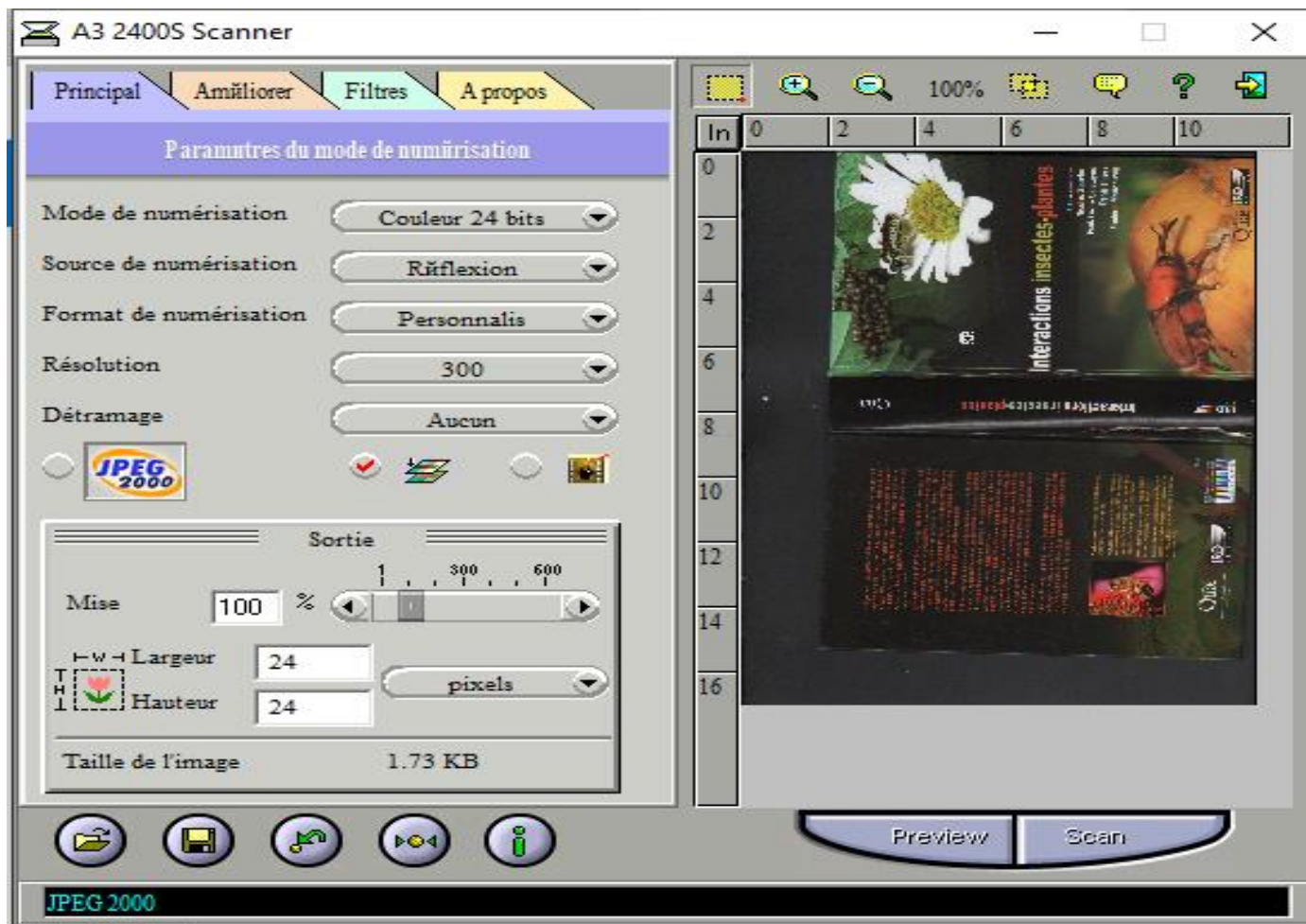


Figure 7 : : Prévisualisation d'un document avant numérisation avec Adobe Photoshop CS2



Figure 8 : Document numérisé avec scanner A3 dur Acrobat Photoshop CS2.

3.3.6 La vérification et finalisation des documents électroniques

La vérification est une étape primordiale dans le processus de numérisation. Cette étape permet de remarquer si :

- ✓ Des pages sont mises en double ou oubliées
- ✓ Des pages orientées dans le mauvais sens de lecture
- ✓ Le document répond à l'original en respectant les couleurs des graphiques, photos, illustrations ou en nuance de gris (photos)
- ✓ La reconnaissance optique des caractères (OCR) de l'ensemble du document est bien faite
- ✓ La saisie des métadonnées pour chaque fichier est respectée

3.3.7 Le reconditionnement du document

Le document une fois numérisé sont soit éliminés après leur massicotage ou remis en état avec une reliure ou encollage au dos. Le document est détruit s'il en existe d'autres exemplaires au niveau du fonds documentaire de l'IRD.

En bref on notera dans ce chapitre que le service IST de l'IRD joue un rôle très important. Il permet aux chercheurs d'enrichir leur recherche en leur donnant des périodiques, ouvrages articles ou de la littérature grise. Le point saillant est le fait que l'IRD s'oriente vers le développement en menant des coopérations avec les pays du Sud.

4 Mesurer l'impact de la mise en ligne et en accès ouvert des publications de l'IRD

Le service IST de l'IRD joue un rôle primordial grâce à sa base bibliographique Horizon Pleins Textes. Il essaye d'améliorer et de multiplier les services documentaires aux chercheurs. Cela passe par une valorisation de leurs travaux. La base Horizon est une base de références comprenant l'ensemble des publications de l'IRD. L'opération Horizon a été lancée le 1^{er} septembre 1986.

Les perspectives sont bien mentionnées dans l'article **de Roux-Fouillet Jean-Paul** (Roux-Fouillet, 1988, p. 286) :

- ✓ Améliorer et à valoriser le traitement documentaire des travaux ORSTOM
- ✓ Diffuser de façon large les résultats de programmes de recherche avec un échange IST sur le réseau francophone de communication

C'est dans cette optique qu'il dira que : « Elle participe à des actions de coopération en matière d'ingénierie documentaire : informatisation des documentations du MESRES camerounais et de ses instituts de recherche notamment. En 1989, l'objectif est d'avoir informatisé l'ensemble des publications ORSTOM depuis 1947 : des actes en ligne seront alors ouverts et l'accès aux documents sera assuré par le recours à un fonds de microfiches (dont la réalisation complète prendra plus de 2 ans) ». **Roux-Fouillet Jean-Paul** (Roux-Fouillet, 1988, p. 287)

Avec les nouvelles technologies le service a connu une évolution. Au début le service utilisait des microfiches. Maintenant grâce au numérique, le travail est devenu plus rapide. Même si l'objectif reste toujours le même.

En 2015, le contrat d'objectifs et de performance de l'IRD propose un indicateur intitulé : Diffuser l'information scientifique au plus grand nombre. Celui-ci évalue la contribution de l'IRD au libreaccès des publications scientifiques en mesurant l'évolution de la part des

publications en libre accès : l'objectif est en effet de réaliser de 2015 à 2020 une progression de 25 % dans la numérisation. En 2016 il y a eu une progression de plus de 8,15 % qu'en 2015. Entre 2015 et 2019 le service de numérisation a réalisé une progression de 25.8 %. Le tableau dessous illustre cette progression, depuis l'année 2010 jusqu'à l'année 2019 :

Valeur de référence 2016 : + 8,15 % /2015

Cible 2020 : + 25 % dans la durée du contrat

Année	PDF en libre accès	PDF en libre accès sur l'année	Progression	Progression 2015-2019
2010	43311			
2011	45811	2500	5.8%	
2012	47711	1900	4.1%	
2013	49041	1330	2.8%	
2014	51288	2247	4.6%	
2015	55619	4331	8.4%	
2016	60157	4538	8.2%	
2017	64765	4608	7.7%	
2018	67236	2471	3.8%	
2019	69974	2738	4.1%	25.8%

Figure 9 : Nombre de PDF mis en accès par l'IRD de 2010 à 2019

Source : Contrat d'objectifs et de performance entre l'État et l'Institut de recherche pour le développement <https://www.ird.fr/>

L'IRD dispose d'une plateforme web qui héberge des bibliothèques numériques créée par la documentation de l'IRD. Elle est une bibliothèque électronique en partenariat avec les pays du Sud. Elle permet ainsi d'avoir accès à plusieurs collections de documents scientifiques produits par les pays en voie de développement.

En attendant que localement des solutions soient mises en place, BEEP³¹ permet d'avoir des solutions face à des problèmes que rencontrent ces pays à savoir : indisponibilité des serveurs, problèmes de connexion, coupures électroniques etc.

Cette coopération permet de valoriser les publications issues de ces pays afin de les rendre accessible aux communautés scientifiques.

³¹ <http://www.beep.ird.fr/cgi-bin/library.cgi>

BEEP produit également des analyses statistiques sur les consultants portant sur le contenu des collections par la mise en place d'outils efficaces. Dans son article (Rossi, 2018) indique que : « En associant, en mode relationnel, la base de données bibliographiques et celle concernant les consultations de ces documents (« fichiers log Apache »), nous pouvons produire des tableaux de bord dynamiques en fonction de plusieurs critères et de plusieurs indicateurs »



Figure 10 : BEEP (Bibliothèque électronique en partenariat)

Source : <http://www.beep.ird.fr/cgi-bin/library.cgi>

4.1 Les statistiques de consultations

Le service IST produit un certain nombre d'indicateurs permettant de mesurer à quel point le fonds numérique de l'IRD joue un rôle majeur dans la diffusion des publications scientifiques, particulièrement dans les pays du Sud. En effet, Ces pays représentent 70 % de l'origine des téléchargements des publications scientifiques IRD.

Dans la figure ci-dessous on voit bien que plus de 70% des PDF téléchargés concernent les pays du Sud. Parmi ces pays on peut citer : la Côte-d'Ivoire, le Cameroun, le Sénégal et le Burkina Faso et le Madagascar. L'exploration de la carte permet en outre de savoir le taux de téléchargements par pays.

Consultations par zone géographique

Avec plus de 8000 pdf téléchargés par jour, le fonds numérique fait l'objet d'une large diffusion dans le monde, et en particulier dans les pays du Sud qui représentent près de 70% des téléchargements.

Retrouvez les taux de téléchargement par pays en 2019 en explorant la carte.

57 pays / 1385916 PDF téléchargés(47%)

Entrer un pays ou une zone géographique

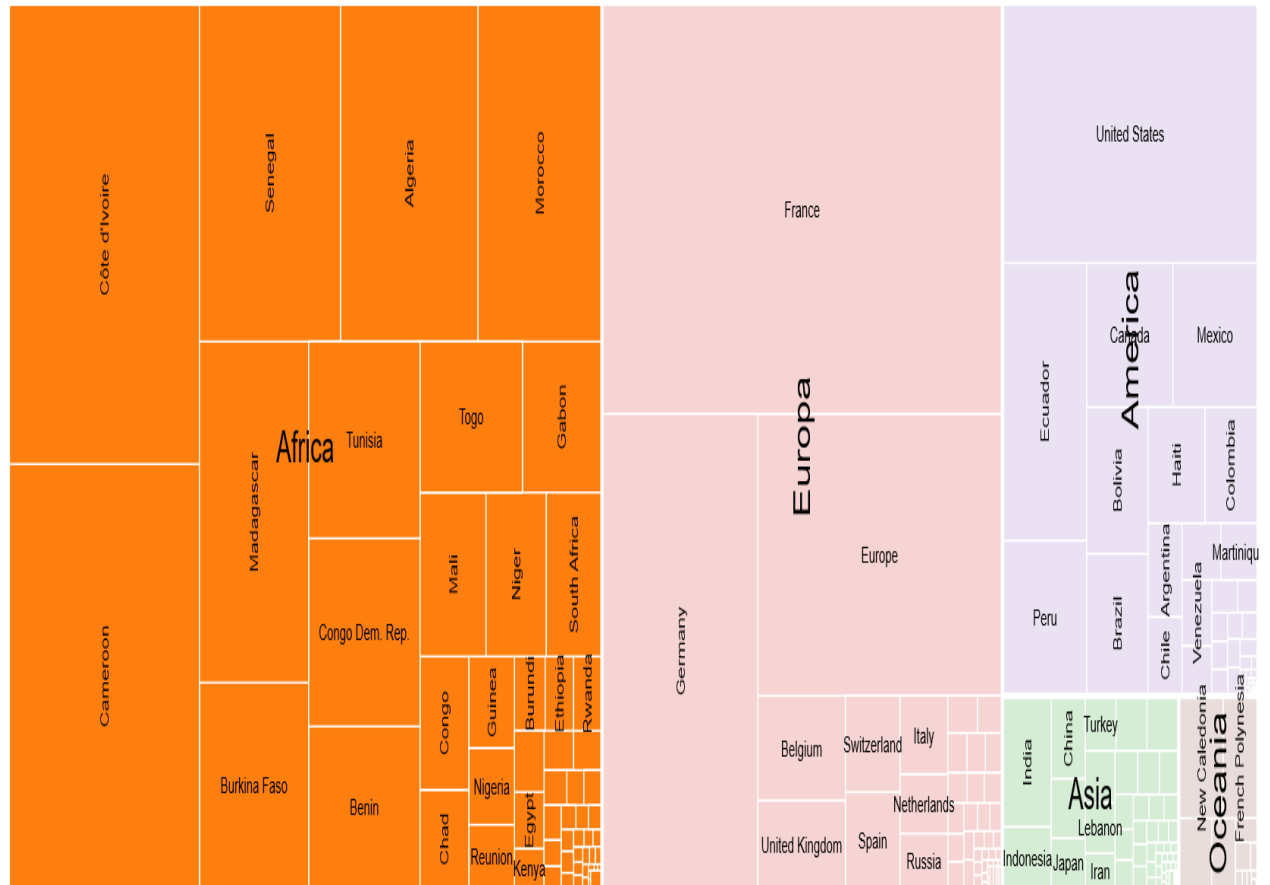


Figure 11 : Consultations zone Afrique

Source : <https://horizon.documentation.ird.fr/>

Sur le tableau de bord des consultations des publications en libre accès, 62 056 fichiers pleins-textes ont été consultés en 2019. Le pourcentage de consultations s'élève à 99.9 % pour la totalité des fichiers pleins textes consultés. La figure ci-dessous montre les différentes valeurs concernant les consultations des publications en libre accès en 2019.

Consultations des publications en libre accès des scientifiques de l'IRD

Tableau de bord "Chapitres"

Tout : présentation des résultats (2019)

Indice	Valeur
Nombre de fichiers "Pleins_Textes" consultés	62056
Nombre de consultations "Pleins_Textes"	2918659
Nombre moyen de consultations par fichier "Pleins_Textes"	47.03
Pourcentage sur la totalité des fichiers "Pleins_Textes" consultés	99.9%
Pourcentage sur la totalité des consultations	100.0%
Facteur D (nombre de documents consultés au moins "n" fois)	680
Top D (680) sur total fichiers consultés	1.1%
Somme des consultations des Top 680	931003
Consultations Top 680 sur total des consultations du chapitre Tout	31.9%

Figure 12 : Tableau de bord des consultations des publications en libre accès des scientifiques de l'IRD en 2019

Source : https://horizon.documentation.ird.fr/exl-doc/pleins_textes/acces/2019.htm

Les indicateurs portant sur les langues des documents et des consultations en 2019 montrent que les documents en langues espagnole et française sont les plus consultés. Le tableau ci-dessous l'illustre bien.

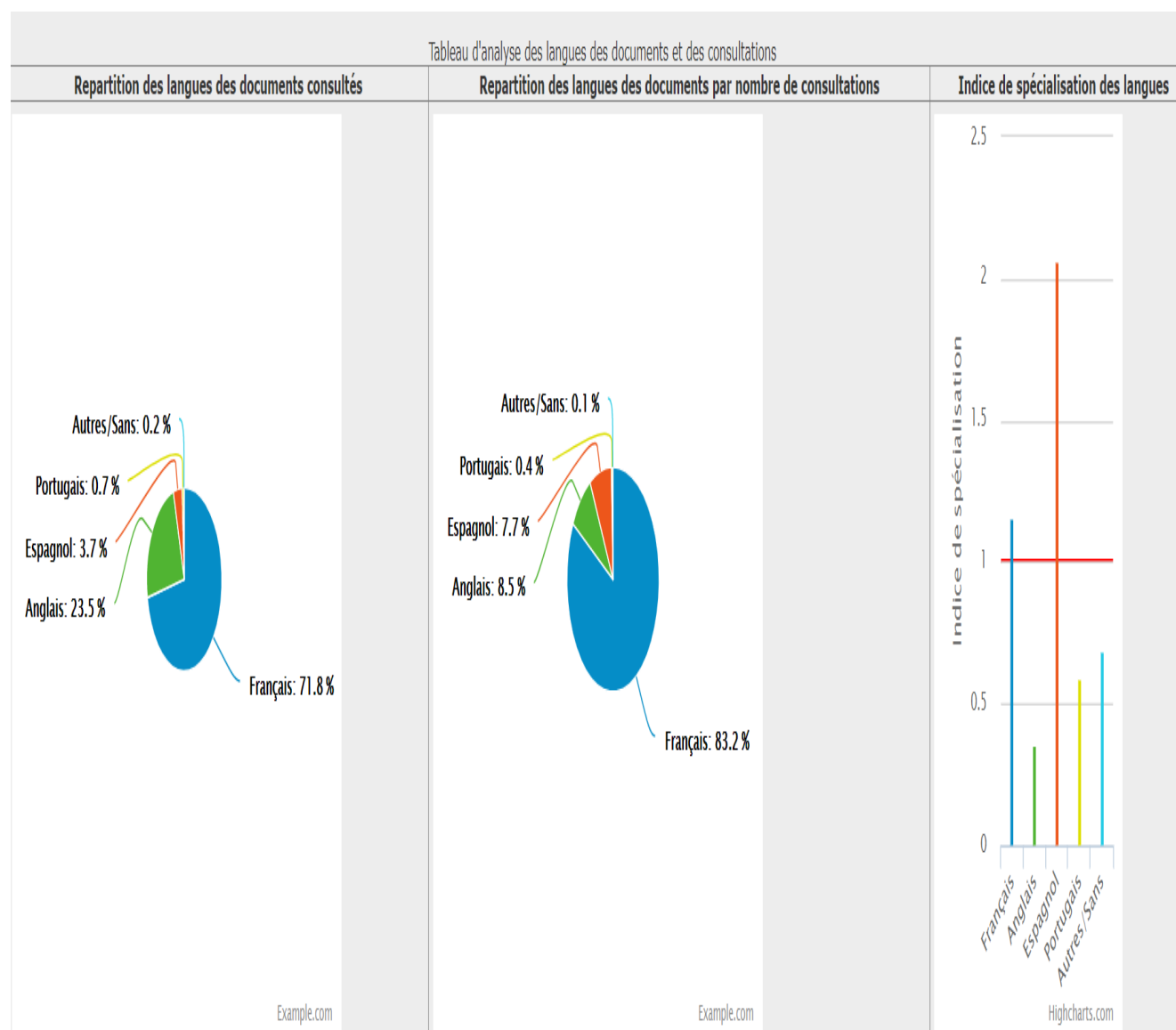


Figure 13 : Tableau d'analyse des langues des documents et des consultations de 2019

Source : https://horizon.documentation.ird.fr/exl-doc/pleins_textes/acces/2019.htm

4.2 Exemples d'indicateurs de comparaison produits par le service de numérisation de l'IRD

Les indicateurs produits par l'IRD permettent de voir de façon particulière pour chaque pays les statistiques lui concernant. Ces statistiques concernent : le domaine d'intérêt pour chaque pays, langues des document, appareils utilisés pour la consultation, marques de téléphones portables utilisés etc. Ici on a des indicateurs de comparaison entre le Burkina Faso, La Côte-d'Ivoire et la France

Les deux premières figures montrent l'intérêt qu'ont le Burkina Faso et la Cote d'Ivoire pour la santé et les sciences humaines et sociales. Le nombre de fichiers consultés au Burkina Faso s'élève à 8343 tandis qu'en Côte-d'Ivoire il est au nombre de 14332.

Burkina Faso : un intérêt pour la Santé et les Sciences humaines et sociales

- 48.475 human accesses
- Nombre de fichiers "Pleins_Textes" consultés : 8.343
- Facteur D : 74
- Moyenne de consultations : 5.8
- Appropriation du pays sur le total des fichiers "Pleins_Textes" consultés : 14.8 %

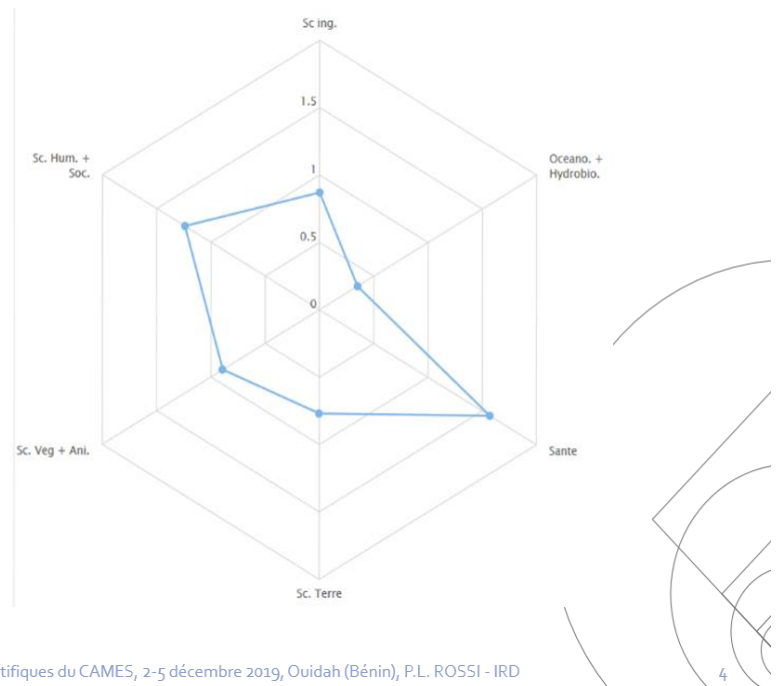
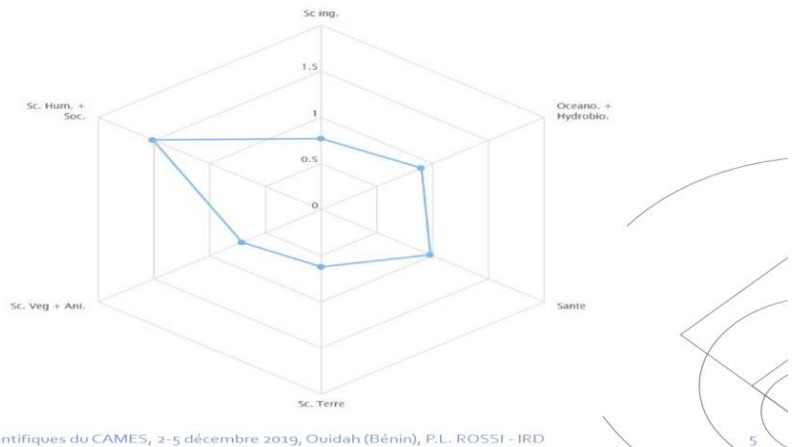


Figure 14 : L'intérêt du Burkina Faso pour la santé et les sciences humaines

Côte d'Ivoire : un intérêt pour les Sciences humaines et sociales

- 183.806 human accesses
- Nombre de fichiers "Pleins_Textes" consultés : 14.322
- Facteur D : 168
- Moyenne de consultations : 12.8
- Appropriation du pays sur le total des fichiers "Pleins_Textes" consultés : 25.4 %

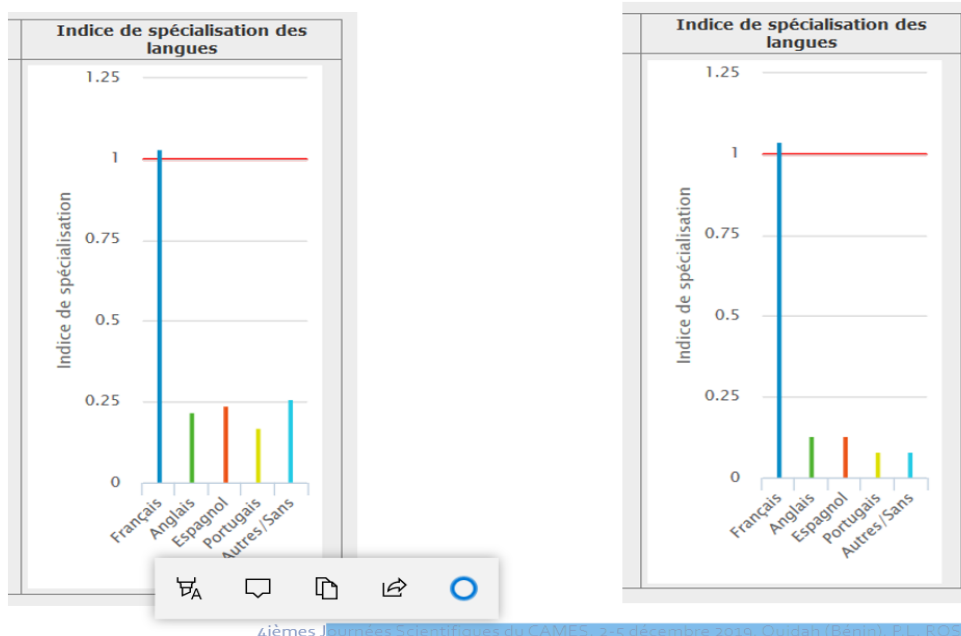


4^{ièmes} Journées Scientifiques du CAMES, 2-5 décembre 2019, Ouidah (Bénin), P.L. ROSSI - IRD

Figure 11 : Côte d'Ivoire : un intérêt pour les Sciences humaines et sociales

La figure ci-dessous montre l'indice de spécialisation des langues des documents consultés. On note que la langue française occupe le premier rang suivi de la langue espagnole et anglaise.

Burkina Faso et Côte d'Ivoire : indice de spécialisation des langues des documents consultés



4^{ièmes} Journées Scientifiques du CAMES, 2-5 décembre 2019, Ouidah (Bénin), P.L. ROSSI - IRD

Figure 15 : Burkina Faso et Côte d'Ivoire : indice de spécialisation des langues des documents consultés

Les deux figures suivantes montrent les appareils et les marques de téléphones utilisés pour la consultation des documents en France, Burkina Faso et la Côte-d'Ivoire.

Burkina Faso, Côte d'Ivoire et France : appareils utilisés pour la consultation des documents

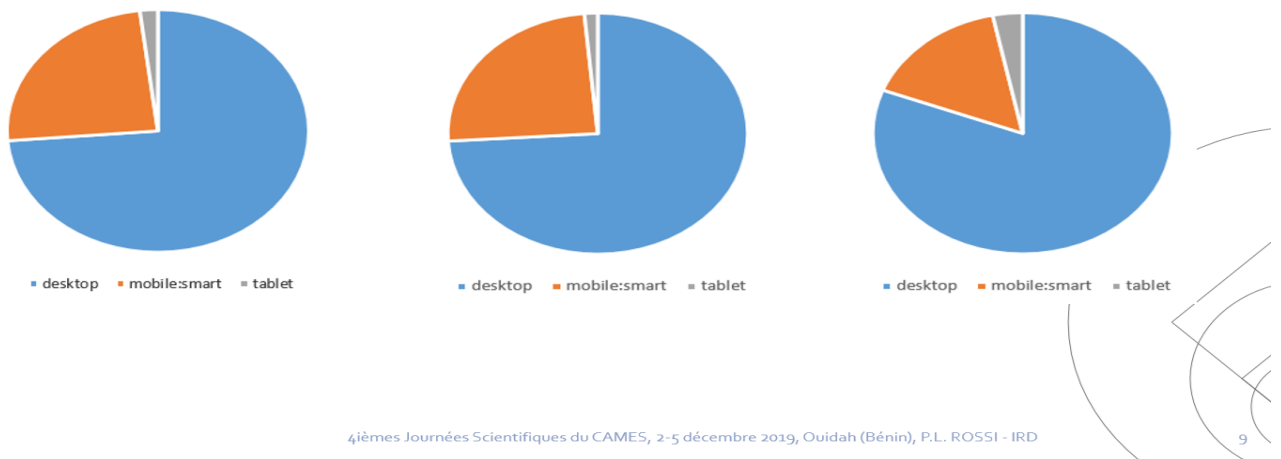


Figure 16 : Burkina Faso, Côte d'Ivoire et France : appareils utilisés pour la consultation des documents

Burkina Faso, Côte d'Ivoire et France : marques des téléphones portables utilisés pour la consultation des documents

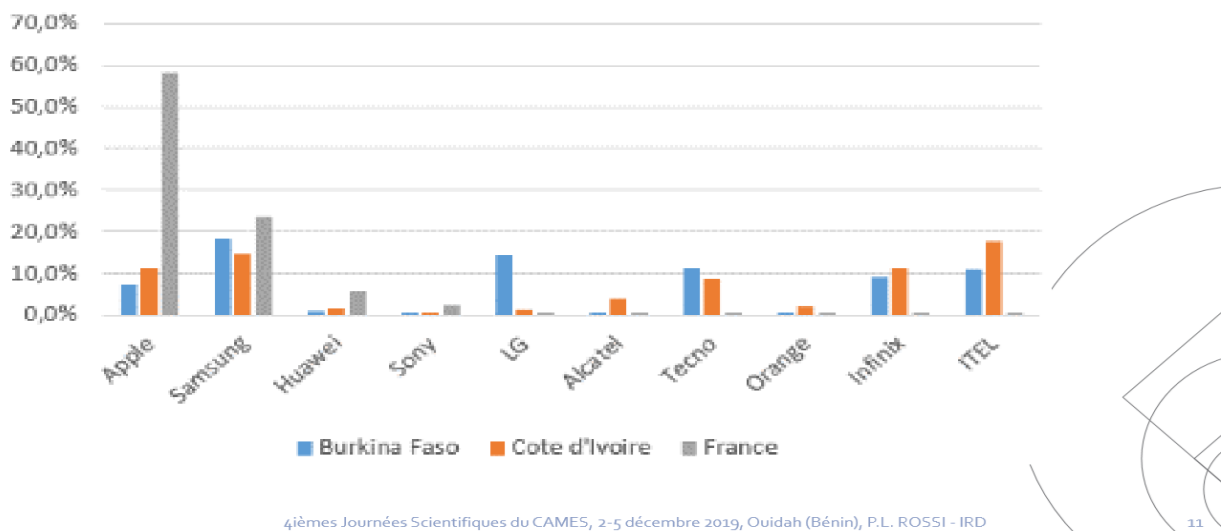


Figure 17 : Burkina Faso, Côte d'Ivoire et France : marques des téléphones portables utilisés pour la consultation des documents

Source : Pier Luigi(Rossi et al., 2019)

4.3 L'enquête qualitative

Parmi les autres outils utilisés pour mesurer l'impact de ses outils, le service IST réalise ponctuellement des enquêtes auprès des scientifiques. En s'appuyant sur l'enquête (voir annexe) réalisée par la responsable du service de documentation, on peut ainsi apprécier l'importance de cette base et son apport aux chercheurs dans la gestion de leurs publications. Beaucoup de chercheurs pensent que la base horizon permet :

- ✓ La visibilité des publications
- ✓ L'évaluation et la centralisation des publications
- ✓ D'avoir une base de référence
- ✓ Un accès aux publications et rapports
- ✓ La valorisation des résultats scientifiques

Quelques intérêts de la base selon les chercheurs :

- ✓ Un accès facile pour les pays du Sud
- ✓ Facile d'utilisation
- ✓ Opérabilité facile entre d'autres bases
- ✓ Durabilité du stockage

En résumé on notera que l'IRD joue un rôle majeur dans la diffusion des publications scientifiques, particulièrement dans les pays du Sud. La mise en ligne et en accès ouvert favorisent un impact positif pour les publications scientifiques.

Conclusion

La numérisation des documents occupe une place très importante à l'IRD. L'institut se veut de lutter contre les difficultés d'accès à la littérature scientifique des pays du Sud. Les Universités de ces pays sont souvent confrontées aux coûts d'abonnements faramineux imposés par les éditeurs qui monopolisent le marché des publications scientifiques.

Le gouvernement français aide beaucoup dans ce sens, en encourageant l'open science afin de mieux valoriser la science qui est une priorité car même les citoyens s'y intéressent également. La libre circulation du savoir est devenue un droit et un besoin naturel.

Le stage m'a permis de mieux connaître le monde professionnel. Il est particulièrement centré sur la numérisation et chaîne de production des documents numériques scientifiques. Il s'est déroulé dans un cadre très agréable et mes collègues en plus d'être gentils, m'ont transmis beaucoup de connaissances et de savoir-faire. Il ressort d'emblée que la numérisation des publications scientifiques nécessite beaucoup de temps et que les tâches ne sont pas simples à réaliser. Elles demandent beaucoup de concentration, de moyens physiques et financiers et d'un savoir-faire.

Le travail au sein du service IST de l'IRD m'a permis dans la pratique de comprendre les attentes et les tâches dans un centre info-documentaire. Le responsable du service de numérisation m'a appris tout le processus de numérisation des documents scientifiques.

Le stage est en parallèle avec mes cours vus sur la gestion de l'information, archivage et record management et de la gestion des documents techniques. Dans ces cours on a eu à voir par exemple : l'archivage et la gestion électronique des documents, la gestion d'un centre info documentaire, la veille etc.

Le service IST fait aussi de la bibliométrie. Elle permet d'évaluer la recherche sur plusieurs niveaux : pays, continent, laboratoire, établissements, chercheur, équipe etc.

Elle se base sur une méthode statistique et mathématiques. Elle est une partie des sciences de l'information. Un article en open access est plus accessible et aura donc plus de chance d'être plus cité.

Plusieurs collections appartiennent à des pays en développement. Les indicateurs produits par l'IRD permettent de voir de façon particulière pour chacun de ces pays les statistiques lui concernant : le domaine d'intérêt pour chaque pays, langues des documents, appareils utilisés pour la consultation, marques de téléphones portables utilisés etc.

Certains manquements au sein du service de numérisation peuvent être signalés comme : certains chercheurs de l'institut qui cèdent leurs droits à des éditeurs privés qui n'acceptent pas souvent une reproduction des documents, la lenteur au niveau du traitement des documents à numériser chez le prestataire, la qualité de certains documents déjà numérisés qui laisse à désirer etc.

Le service fait tout son possible pour combler ces différents manquements afin de continuer à répondre à une demande forte particulièrement des pays du Sud.

Le stage m'a permis de mieux comprendre l'importance d'une publication scientifique. J'ai eu l'occasion d'apprendre beaucoup de choses sur la recherche scientifique et sur la vie professionnelle.

Bibliographie

- Barts, N., & Cavet, D. (2014). *Suivi des publications IRD : Rapport annuel portant sur l'année 2013 et les évolutions entre 2004 et 2013* (Centre IRD de Bondy ; p. 53 p. multigr.). IRD. <http://www.documentation.ird.fr/hor/fdi:010064292>
- Bordignon, F., & Andro, M. (2016). Impact de l'Open Access sur les citations : Une étude de cas. *I2D Information, donnees documents, Volume 53(3)*, 70-79.
- Cole, F. J., & Eales, N. B. (1917). *The history of comparative anatomy : Part 1. - a statistical analysis of the literature*. Science Progress.
- DIST. (2017). *Livre blanc—Une Science ouverte dans une République numérique—Guide stratégique : Études et propositions en vue de l'application de la loi*. OpenEdition Press. <http://books.openedition.org/oep/1706>
- Lawrence, S. (2001). Free online availability substantially increases a paper's impact. *Nature*, 1-2. <https://doi.org/10.1038/nature28042>
- Lubkov, M., & Roux-Fouillet, J.-P. (1990). L'informatique documentaire à l'ORSTOM. *Archimag*, 35, 34-36.
- Office de la Recherche Scientifique et Technique Outre-Mer : Organisation—Activités : 1944-1955*. (s. d.). 188.
- Rossi, P. L. (2005). *La numérisation des documents au format pdf et la réalisation d'une bibliothèque électronique avec le logiciel Greenstone : La bibliothèque d'information scientifique et technique du Burkina Faso*. 8.
- Rossi, P. L. (2018a). Indicateurs de consultation, indicateurs stratégiques : Leur production à partir de l'analyse des consultations des documents d'une archive en libre accès. In N. Ibnlkhayat, A. A. Bachr, A. Benchakroun, & O. Roudiès (Éds.), *Le libre accès à la science : Fondements, enjeux et dynamiques : Actes du 3e colloque international sur*

- l'Open Access* (Centre IRD de Bondy ; Numéro 1, p. 222-237). Centre National de Documentation. <http://www.documentation.ird.fr/hor/fdi:010074292>
- Rossi, P. L. (2018b). *Numérisation, bibliothèques électroniques et libre accès : Entre renforcement de capacités et perspectives en Afrique francophone*. 13.
- Rossi, P. L., Journées Scientifiques du CAMES, 4., Ouidah (BEN), 2019/12/02-05, & Journées Scientifiques du CAMES, 4., Ouidah (BEN), 2019/12/02-05. (2019). *Statistiques de consultations d'une collection de documents scientifiques en libre accès : Comparaison des comportements des internautes du Burkina Faso et de la Côte d'Ivoire [résumé]* (Centre IRD de Bondy).
<http://www.documentation.ird.fr/hor/fdi:010078241>
- Roux-Fouillet, J.-P. (1988). Horizon : Base bibliographique ORSTOM : présentation. In P. Séchet (Éd.), *Séminfor 1, premier séminaire informatique de l'ORSTOM : bases de données et systèmes d'information : Quelles méthodes ?* (Centre IRD de Bondy ; p. 285-296). ORSTOM. <http://www.documentation.ird.fr/hor/fdi:26249>
- Suber, P. (2016). *Qu'est-ce que l'accès ouvert ?* OpenEdition Press.

Annexes

Annexe 1 : Organigramme de l'IST

Annexes 2 : Dessins open access vert et gold

Annexe 3 : Enquête du service IST sur la base Horizon auprès des chercheurs de l'IRD

Organigramme 2018 du service IST

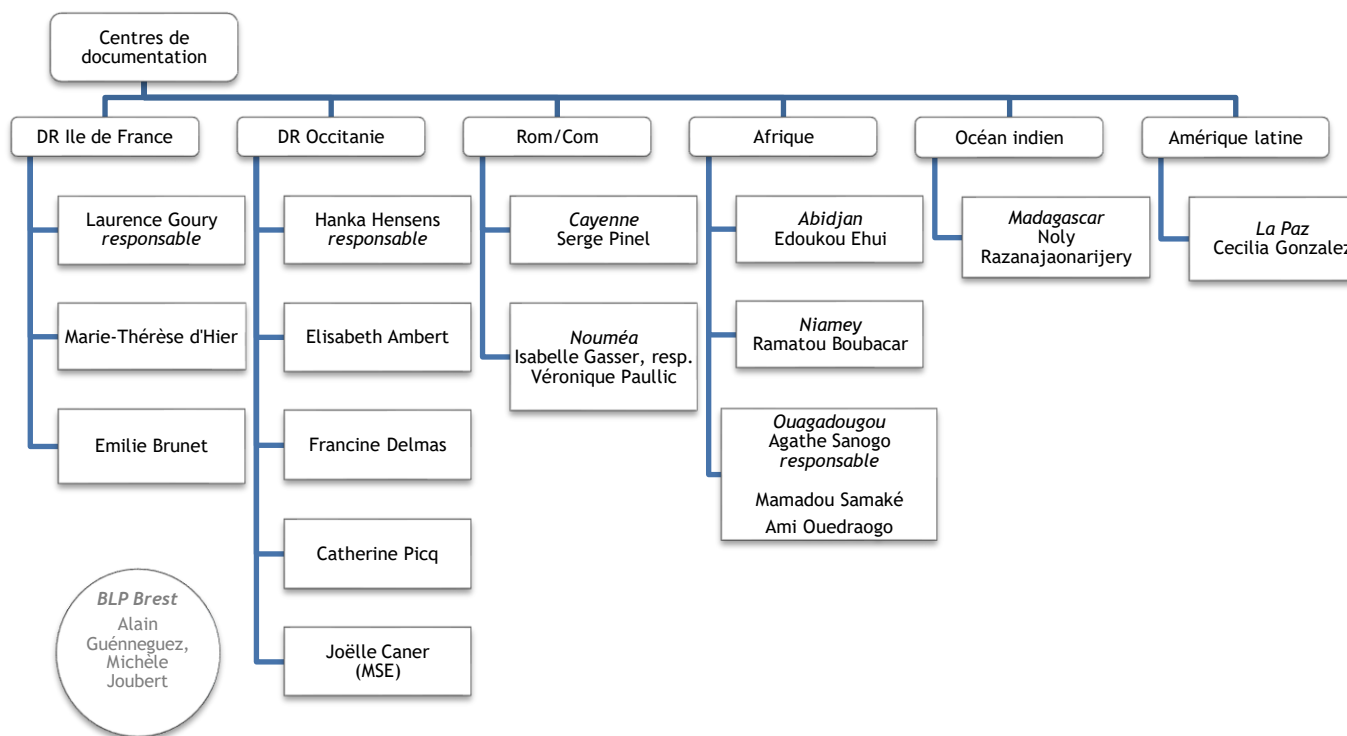
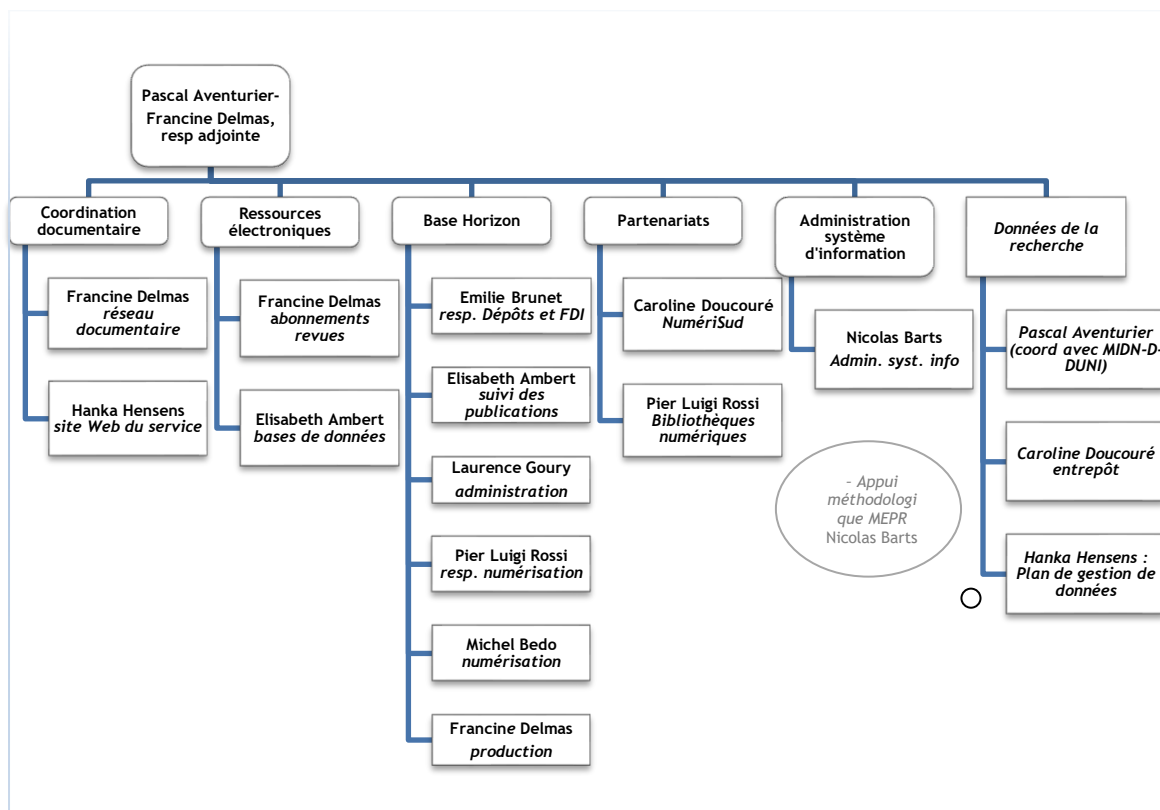




Figure 1 : La voie verte d'open access

Source : <https://scienceouverte.couperin.org/la-voie-verte-2/>



Figure 2 : La voie dorée d'open access

Source : <https://scienceouverte.couperin.org/la-voie-doree-2/>

Questionnaire sur les motivations du dépôt dans Horizon				
Réponses suite au mail du 04/06/2020				
46 réponses au 18/06/2019	Quelles motivations pour déposer dans Horizon?	Quels avantages?	Dépôt dans Hal ou ailleurs	Commentaires
Chercheur 1	pas prise en compte uniquement des publiés dans revues internationales, mais aussi littérature grise, fait un lien avec Orid	solution optimale en particulier pour diffuser auprès des partenaires	non, par manque de temps et actuelle multiplication des bases	
Chercheur 2	rendre accessibles mes publications	visibilité et durabilité du stockage	non	rien de spécial, Horizon est un dispositif très utile
Chercheur 3	Rendre les publications le plus facilement possible accessibles et visibles. Répondre à l'invitation de l'IRD et de la direction du laboratoire.	J'aime particulièrement dans Horizon la possibilité de mettre en ligne la littérature grise. C'est un autre type de documents très précieux quand on veut revisiter un terrain ou dans le contexte des observatoires notamment.	Oui. Je dépose sur HAL ce qui peut l'être en utilisant l'identifiant horizon.	La chaîne d'affichage sur HAL -> affichage sur le web labo LPED est automatique.
Chercheur 4	Regrouper quelque part toutes les publications IRD concernant les Observatoires du Sénégal. visibilité pour mon labo,	diffusion et visibilité auprès des collègues, en particulier francophones	non, parce que je croyais que dépôt sur Horizon impliquait dépôt sur HAL	excellente Base biblio !!
Chercheur 5	prémiation, visibilité, bancarisation travaux IRD, base unique pour les partenaires africains		HAL: non (très rarement) / Researchgate (souvent), OIRD (rarement)	Ouvrir le dépôt a des Sud non IRD mais parainné par un IRD
Chercheur 6	Cette base est la mémoire historique de notre institut en particulier pour la science du sol. C'est aussi le support de conservation de mes propres documents. - Elle est la résultante d'années de sélection, conservation - Lien d'accès pour citer une référence bibliographique non publiée dans un article de revue à comité de lecture - Transfert rapide de la doc dans ORCID - Consulter la production de mes collègues	Facile d'accès, complète - Mise à disposition de TOUS les documents mis en ligne, téléchargeables et accessibles par tout le monde, en particulier le Sud - Lien d'accès pour citer une référence bibliographique non publiée dans un article de revue à comité de lecture - Transfert rapide de la doc dans ORCID - Consulter la production de mes collègues	ORCID	TRES BIEN travail à poursuivre.
Chercheur 7	Bon endroit pour rendre mes publications accessibles	Accès facile notamment pour les partenaires du Sud et pour les étudiants	Régulièrement dans HAL, notamment les traductions françaises d'articles publiés en anglais (en plus de la version anglaise)	Rien de spécial
Chercheur 8	je dépose car on m'en a demandé, mais aussi car je trouve que c'est une sauvegarde très commode et efficace.	En outre, cela permet de rendre facilement visible ma production en cas d'évaluation, et cela permet aussi de voir facilement la production de mes collègues	J'évite de déposer dans Research Gate, mais ils en ont pas mal tout de même. J'ai beaucoup de mal à déposer dans hal que je trouve moins ergonomique et moins convivial. En archéologie, où l'on a beaucoup de figures, il est très difficile de fabriquer un texte libre de droit avec les figures incluses... Si je devais reprendre tout ma biblio pour la mettre en format word-PDF avec les images, ce serait beaucoup de temps passé pour rien. Une pure répétition du travail.	continue !
Chercheur 9	cela m'était utile pour mes rapport d'activités biannuels (je faisais des copier-coller). Mais depuis un an ou deux il a été rajouté des logo (doi, I) si bien que le copier-coller rajoute plein d'images à effacer manuellement, ce qui perd de l'intérêt par rapport aux années précédentes, le cherche parfois de la littérature grise IRDienne sur Horizon.	Je ne dépose pas dans HAL directement, mais la bibliothèque d'AMAP le fait. Je dépose aussi à la bibliothèque du labo, d'ailleurs je fais en général un seul mail (biblio AMAP / dépôt horizon) pour être sûr de bien mettre mes articles dans ces deux bases.	Je dépose dans ResearchGate (en libre quand mes articles sont en open access, en copie privée les autres cas).	
Chercheur 10	Registre des productions IRD, listing "valide" sur la forme et le type de publiés dans la base horizon, intérêt pour l'évaluation IRD.	Réutilisation pour alimenter les BD de l'UMR GREED pour évaluations HCERES, gestions de projets ou demandes des représentations IRD, consultation par nos partenaires et étudiants au Sud	Je n'ai pas déposé ailleurs, HAL ou autres (1) par manque de connaissance sur les avantages et inconvénients, (2) par manque de temps. Consistent du manque de visibilité extérieure de mes travaux, malgré sollicitations presque journalières de ACADEMIA et ResearchGate qui encombrèrent ma boîte mail.	Question: est ce qu'une référence Horizon avec son numéro de code peut remplacer un DOI?
Chercheur 11	Sur le papier c'est très bien que toute la production de l'IRD soit accessible en un seul endroit. Le problème c'est que les pdf originaux ne sont pas vraiment accessibles.	Certains organismes le font pour moi ...		Mais il y a vrai dire l'ensemble du système est très peu satisfaisant. Les articles sont déposés à plein d'endroits (sans parler du recueil de la production scientifique de mon labo en vu de l'HCERES) ... mais au final les chercheurs ont tendance à utiliser researchgate et google scholar pour rechercher la biblio et récupérer directement les pdf, et qui sont des systèmes très efficaces (mais privés...) ! Et évidemment, tout ça ne remplace pas un bon système de publication, non tenu par le privé, et coûtant bcp moins à la recherche publique.
Chercheur 12	Pour les chercheurs IRD, c'est une obligation (avec laquelle je suis en total accord).	Ça permet de garder une trace. Par ailleurs, ça permet également de garder un accès gratuit à nos travaux de recherche.	Oui dans HAL, mais pas ailleurs (j'ai un compte Researchgate, mais je n'ai pas l'habitude de mettre les documents en libre accès dessus).	Il faut continuer à maintenir la base de données Horizon. Grâce à elle, j'ai accès à des documents anciens qui sont une source d'informations précieuses auxquelles on n'aurait plus accès si ces documents n'avaient pas été sauvegardés dans la base de données. Bravo pour le travail que vous faites.
Chercheur 13	Démarche réglementaire nécessaire pour l'évaluation, centralisation des publications d'un auteur	Accès gratuit, pérenne, exhaustif (pour ceux qui envoient...), utile au chercheur lui-même pour retrouver ses propres publications	Non	Créer des passerelles automatiques de référencement des publications (Horizon->UMR) pour les chercheurs IRD des UMR. Dans le cas de l'UMR Ceped, nous devons aussi déposer nos publications sur le site du Ceped avec une procédure différente de celle de Horizon.
Chercheur 14	par ordre... : 1. Affichage pour les rapports d'activité / habitude de répondre aux demandes de l'IRD... 2. Affichage pour collègues et extérieurs 3. Archivage (facile de retrouver les textes et l'apprécier la mise aux normes de présentation)	Cela oblige à organiser les publications et les rend disponibles par la suite	Je devrais déposer dans HAL... j'attends que l'IRD verse Horizon dans Hal...	Etablir des passerelles avec HAL
Chercheur 15	je dépose mes publications dans Horizon, car il me semble que c'est indispensable pour les évaluations et pour l'obtention des financements à l'IRD, mais aussi il me semble essentiel de pouvoir diffuser mes travaux au sein de l'institut.	Je n'ai rien déposé sur HAL pour l'instant. Je ne retrouvais plus mes codes d'accès. D'après ce que j'ai compris, le référencement passe maintenant directement de la base Horizon à HAL. Je suis sur Orid, ResearchID et Publine. J'ai déposé quelques travaux sur Academia.edu et ResearchGate, mais sur vos conseils, je ne dépose plus rien (mais je crois que des co-auteurs l'ont fait), je n'indique plus que les références. J'ai également déposé des articles sur le site du GEMESO, un groupe de chercheurs travaillant sur le Mexique.		Ce serait bien que vous nous envoyiez un petit paragraphe sur ce que nous devrions faire ou ne pas faire sur des sites tels que Academia.edu et pourquoi.
Chercheur 16	Horizon est la base de données de référence pour les évaluations IRD, individuelles et collectives, de l'activité de recherche. Il est donc important qu'elle soit complète et à jour.	Les textes sont durablement dans la base de donnée et on peut y faire référence dans différents documents avec l'assurance que les liens hypertexte ne changeront pas. Elle permet de visualiser l'ensemble de la production scientifique d'une personne ou d'une équipe.	Je ne dépose pas sur HAL car le processus est plus long et contraignant que sur Horizon et je sais que les infos sont disponibles sur Horizon. Je dépose aussi sur ResearchGate car cette base de données est très consultée et globale. Une partie de mes publications sont aussi disponibles sur la base Agora du CIRAD car j'ai été détaché plusieurs années au CIRAD et que le dépôt sur Agora est très facile. Mes articles dans des revues de rang A sont aussi référencés sur les bases de données Scopus et Orid, etc., sans que je n'aie rien à faire qui est bien pratique.	Les points essentiels : - La durabilité de la base de donnée et la stabilité des liens hypertexte des références. - La simplicité de dépôt des articles. - L'interopérabilité de bases de données pour mettre à jour différentes bases en ligne à partir de celle qui est la plus complète - Horizon.
Chercheur 17	Ma première motivation a été la visibilité à l'intérieur de l'IRD (notamment pour évaluations, concours, avancement) car j'avais l'impression que la base Horizons n'apparaissait pas quand on tapait mon nom sur Google	Je peux retrouver rapidement les références de mes publications. D'autres bases de données y récupèrent des références pour les intégrer dans leurs propres bases (ex. Academia, ResearchGate)	Après avoir suivi le mois dernier une formation en ligne de HAL pour apprendre à déposer j'ai appris qu'on pouvait faire des transferts depuis Horizons et j'ai ainsi déposé tous mes articles. Par contre ça n'a pas fonctionné pour les chapitres d'ouvrages et je devrai donc rentrer les données moi-même...	Rendre possible le transfert de chapitres d'ouvrages de la base Horizons vers HAL...
Chercheur 18	Parce que c'est ce que me demande l'IRD	Disponibilité supplémentaire pour un public encore élargi	Je dépose mes articles dans mon site web dès qu'ils ont une année d'exploitation par les maisons d'édition et à la demande dans Research Gate	
Chercheur 19	La base Horizon est une base que j'apprécie beaucoup, qui a été pionnière (je l'ai beaucoup utilisée déjà en tant qu'étudiante dans les années 1996-2000 et c'était un progrès gigantesque par rapport aux photocopies) et je souhaite en tant que chercheur IRD y contribuer autant que possible, en pensant à nos étudiants, particulièrement aux étudiants du Sud.	Le dépôt y est déjà plus facile que dans HAL et les références sont plus précises. Il permet de connaître le travail de nos collègues IRD. Il est possible de trouver des anciens documents et rapports, ce qui est très utile. Un autre avantage est qu'elle contient des ouvrages, et pas que des articles.	Je dépose également dans HAL et dans ResearchGate quand j'en ai l'autorisation.	Merci pour votre travail
Chercheur 20	L'importance d'avoir une visibilité de la production bibliographique propre à l'IRD. Je recoure assez souvent au site horizon pour consulter des références pour la recherche actuelle bien sûr mais aussi pour comprendre l'évolution des approches scientifiques autour de tel ou tel thème. Pour moi c'est un instrument IRD indispensable de visibilité, encore plus pour nos partenaires étrangers pour lesquels l'UMR/Initiation reste souvent floue et encore plus si on ajoute une couche de COMMUNICATION.	La relative interopérabilité entre bases de données, HAL, Horizon et les bases de recherche type web of knowledge, scielo...	Oui pour certains publiés qui ne sont pas forcément sur Horizon, des productions antérieures.	Les questions de passage sur EndNote ou autre logiciel qui se sont améliorées mais où tout n'est pas encore possible.
Chercheur 21	Autre élément, la possibilité d'avoir accès à des publications, rapports... d'avant le tout digital, et, en tant que chercheur pas trop loin de la retraite, d'être sûr que le non digital n'est pas perdu. C'est aussi un instrument d'usage quelque peu obligatoire pour les évaluations. pour que toutes ces publications soient accessibles sur le net pour les chercheurs et la base Horizon	Je pense que cela permet au chercheur que je suis de mieux se faire connaître et de faire connaître la diversité de mes champs de recherche	Ce n'est pas exactement du dépôt mais mes publiés sont sur les CV en ligne du CNRS (base de données Lattes: http://lattes.cnrs.fr/) excellent instrument, démocratique, ouvert, un peu laborieux comme investissement au départ mais qui permet d'avoir la cartographie scientifique de ce qui se fait dans un pays ou par les chercheurs de ce pays. On aurait pu se doter d'un tel instrument en France !	
Chercheur 21	pour que toutes ces publications soient accessibles sur le net pour les chercheurs et la base Horizon	je pense que cela permet au chercheur que je suis de mieux se faire connaître et de faire connaître la diversité de mes champs de recherche	oui dans Hal	

	On simplifie ainsi la découverte de certains articles que n'auraient pas été référencés. outil institutionnel qui était déjà en cours avant les grands sites et/ou réseaux de citation scientifiques			dépôt dans HAL effectué par le labo; je m'occupe de tenir mes publications et quelques actions de recherche (conf, talk invited, etc...) dans ResearchGate	Le service devrait proposer une maj des publications pour le fichier CV des évaluations des personnes.
Chercheur 22	La mise à jour de ma liste de publications pour son utilisation dans mon rapport d'activités.	accès interne à notre production scientifique	Le dépôt sur Horizon est simple et il me facilite la tâche lors de la soumission de mon dossier sur le portail du système d'information.	Oui ça m'arrive, mais ce n'est pas pratique. Le système HAL est obscur.	Non comment. Merci pour le travail effectué sur Horizon!
Chercheur 23				J'ai essayé de comprendre, mais ce n'est toujours pas clair pourquoi des publications peuvent être déposées dans le système et d'autres non. Je dépose aussi mes articles sur LinkedIn et ResearchGate.	
Chercheur 24	***plus de diffusion et en accès libre, et je crois me souvenir que l'on m'avait dit qu'il fallait le faire pour les statutaires, leur évaluation et les dossiers des labos	***accès à un large public et entré sur le développement	***oui		***je suis parfois un peu perdu entre HAL, Horizon et les demandes de mon Labo de déposer aussi nos articles dans notre Zotero interne, donc pour chaque article cela me demande plusieurs envois et dépôts. ce qui serait aidant pour tous je crois c'est de centraliser cela une seule fois, surtout pour nos articles qui sont dans des revues référencées et donc facile à trouver pour les autres
Chercheur 25	Actualisation de mes principales productions sur notre site institutionnel. Horizon est le lien naturel en interne (comme lors des évaluations par ex.).	Lisibilité	Researchgate (HAL : pas encore mais je vais m'y tenir).		RAS - vous êtes utiles ! Continuez !
Chercheur 26	mettre à jour la base interne IRD qui peut être consultée par l'administration et les commissions assurer une large visibilité des articles car Horizon est indexé dans Google Scholar et tout ce qui y est entré est rapidement visible sur Scholar		Plus maintenant du fait de la passerelle.		Créer un lien avec ResearchGate pour que RG appelle la base Horizon, comme c'est fait pour les data
Chercheur 27	Obligation statutaire non ?	Non pas me faire vivre? ... Mais c'est aussi pour rendre visible ma production aux autres, au public et ça me sert moi-même à la classer.	non mais ça fait toujours pas arriver dans google scholar et autres.		J'aimerais que vous nous transmettiez les fichiers pdf que vous produisez à partir de publication papier, pour nos archives
Chercheur 28	C'est le minimum vu-à-vis de notre administration et entre nous c'est une obligation je crois bien.	Bien pratique pour les rapports d'activités Toutefois, je suis aussi content du travail fait par les gestionnaires de la base HORIZON qui valorisent bien son contenu.	Non je ne crois pas mais je suis actif sur l'internet : blogs chez Futura " Climat o " et chez Mediapart		Bravo et continuez SVP !
Chercheur 29	Mise en Visibilité de mes recherches + le fait qu'officiellement la base horizon tient de référence lors des évaluations bi annuelles.	Mise en visibilité de mes recherches	Non, pour HAL: besoin de "déformater" les papiers acceptés dans des revues et être en accès non libre (ou d'utiliser une ancienne version du manuscrit qu'il n'est pas toujours facile de retrouver notamment dans le cas où je ne suis pas l'auteur principal de la publication)		Je vous remercie pour votre appui en ce qui concerne le dépôt des publications que vous rendez extrêmement facile.
Chercheur 30	Je ne sais pas si on peut parler de motivation, mais sauf erreur, il nous est demandé de déposer dans Horizon pour que nos publications soient prises en compte pour nos rapports d'activités etc... En parallèle de cela, ça permet de rassembler nos différents travaux dans un même endroit, de leur conférer « l'identité IRD », et de contribuer par ailleurs à la remontée des activités de notre institut.	Pour la littérature grise, ça permet de la rendre plus accessible d'assurer sa pérennité dans le temps.	Dans HAL, oui ponctuellement. Les articles co-écrits avec des collègues de l'Ifremer se retrouvent automatiquement déposés dans Archiver. J'utilise également ResearchGate.		J'ai appris sauf erreur que le dépôt dans Horizon enclenchera automatiquement le dépôt dans HAL, c'est une très bonne chose, merci ! Dû en est-on de la possibilité d'attribuer un doi à des rapports, jeux de données etc ? La question s'était posée pour notre eBook Sponges of Polymia. Il y a 3 ans, les choses étaient en train de se mettre en place, mais depuis je ne sais pas où cela en est.
Chercheur 31	Garder trace	Permettre accès lecteurs IRD et si possible externes			Renvoi automatique du rapport d'activité, point qui mériterait d'ailleurs d'être formalisé et développé pour éviter les duplications
Chercheur 32	Accroître la visibilité institutionnelle de l'IRD	Les dépôts sont faits de manière professionnelle par votre service.	Oui, dans HAL. Cela nous est demandé par le laboratoire et je pense que cela accroît la diffusion scientifique des travaux (plus qu'Horizon, qui est à mon avis surtout institutionnel).		Je me demande depuis plusieurs années s'il est utile de maintenir Horizon et HAL-IRD en même temps et si le travail de votre service ne pourrait pas être investi dans HAL-IRD, que les laboratoires nous demandent d'alimenter et qui a une plus grande visibilité scientifique car il est plus général.
Chercheur 33	C'est pour moi une question de déontologie professionnelle, que je suis depuis mon recrutement à FORSTOM/IRD: il me semble normal de déposer dans la base de l'IRD - mon employeur - les résultats des recherches (sous forme de publications et communications) que l'IRD a contribué à financer (et ce de manière largement majoritairement si l'on prend en compte le salaire en plus des budgets de fonctionnement), sans parler du soutien logistique et institutionnel.	Cela assure (i) une meilleure visibilité à mes publications, (ii) un stockage sûr de mes productions, (iii) une mise à jour en continu de la liste de mes publications, que je peux réutiliser pour divers dossiers (évaluations, bilans, etc.)	Oui, mais depuis deux ans seulement, et pas nécessairement les textes intégraux. OUI: Academia et Research Gate, mais je ne fais de dépôt régulier de la liste de mes publications, que je peux réutiliser pour divers dossiers (évaluations, bilans, etc.)		Le versement automatique des dépôts dans la base Horizon vers HAL est une excellente chose. Merci pour le travail de toute l'équipe,
Chercheur 34	« contribuer à la "bibliothèque de publication" de l'IRD, avoir un lien où se trouve l'ensemble de mes publications	archiver et rendre accessible mes publications	Dans HAL et avant d'être à l'IRD dans Academia et researchgate		Intégrer les publications avant l'entrée de l'agent à l'IRD + établir des liens automatiques HAL-Horizon (si ce n'est pas déjà à l'air?)
Chercheur 35	Obéissance à la règle	Visibilité de l'IRD	non		néant
Chercheur 36	Il s'agit d'une base bibliographique de l'IRD, j'y dépose donc naturellement mes publications.	1) une accessibilité plus facile pour les collègues et étudiants du Sud, qui souvent n'ont pas accès aux ressources en ligne (après tout, les recherches sont la plupart du temps relatives au Sud); 2) je pense de plus que cela facilite, pour l'IRD, l'obtention de ses statistiques (nombre de publications, IF, etc.)	HAL non, j'avoue pas savoir comment faire. Je dépose également à la bibliothèque de mon LMR (ISE-M).		Serai-il possible que, dans les ouvrages dont on est éditeur, les chapitres dont on n'est pas auteur n'apparaissent pas dans notre bibliographie ?
Chercheur 37	visibilité au sein de l'IRD et à l'extérieur	ben avec HAL	Oui HAL, ResearchGate, Academia		merci pour votre soutien !!
Chercheur 38	La possibilité de rendre accessible mes publications	Pas d'avantage particulier me concernant, mais l'accès libre pour les collègues notamment ceux du Sud	Oui, et je souhaiterais à cela était possible qu'il y ait des "ponts" entre les deux afin d'éviter de devoir déposer deux fois.		Un dépôt "unique" HAL et Horizon...
Chercheur 39	la valorisation des résultats scientifiques pour plus de visibilité: l'IRD est réputé à l'international alors le fait de publier avec l'IRD valide certains avantages.	Visibilité et validation scientifique	Oui, je dépose sur Hal, sur Researchgate, sur googlescholar, pubmons ...etc.		
Chercheur 40	Deux motivations principales : faire connaître mes publications, la recherche sert à partager les publiés ! Donner de la visibilité au laboratoire et à l'IRD et faciliter la gestion des publiés pour les rapports, etc.	C'est un système de l'institution et c'est important que l'institution ait une vitrine des publications.	Toutes les publications sont déposées dans HAL, car c'est la référence aujourd'hui au niveau universitaire. C'est d'ailleurs obligatoire pour que l'université nous attribue des budgets.		Je pense que Horizon et HAL sont interconnectés ? En tout cas il le faudrait.
Chercheur 41	L'évaluation biennale me fait penser tous les 2 ans à déposer mes publications dans Horizon	Je ne vois pas trop d'avantages par rapport à ORCID ou ResearchID.	Non		RAS
Chercheur 42	L'archivage des publications et documents de travail (pour les autres chercheurs et pour soi-même)	Conservation des documents	Oui, dans HAL, en général via Horizon		Continuez. Le service est excellent et très important.
Chercheur 43	Que mes travaux IRD soient visibles sur le portail institutionnel	Possibilité de les retrouver facilement - Mise à disposition pour les autres chercheurs - Documentation de l'ensemble des travaux du chercheur - Publicité des travaux	HAL (mais maintenant le lien est fait directement, c'est super) + bases de données (Zenodo + Nexa)		- On aimerait bien aussi pouvoir archiver de la même manière, dans un autre service dédié (comme Indigo pour les photos), les autres matériaux des travaux de recherche, (questionnaires, méthodes, archives de terrain, banques de données, etc.), sous forme scannée pour les documents papier et électronique pour les données.
Chercheur 44	Mémoire indispensable de la valorisation + nécessaire pour la visibilité + nécessité pour les évaluations	On peut déposer tous types de travaux. La recherche est rapide et fonctionnelle. On peut facilement partager les références. L'avantage est aussi l'aide reçue pour le service en charge de gérer les dépôts (efficacité, amabilité, ponctualité)	Le service Horizon effectue le dépôt pour moi, car je n'ai pas eu le temps de suivre la formation HAL pour effectuer les dépôts moi-même.		Intéressé Horizon/Hal biennames. Pour valoriser l'IRD (en le service) je pense qu'il serait bon d'envoyer à tous une fois par mois une brève synthèse chiffrée des publications déposées (par département).
Chercheur 45	Avoir une synthèse disponible en ligne de l'ensemble de mes publications qui permet notamment une récupération facile dans les bases bibliographiques comme Zotero.	J'y vois un lieu de stockage permanent et des notices produites par des professionnels, avec l'avantage d'y avoir les documents issus de la littérature dite "grise", p. ex. rapports et communications aux conférences.	Je déposais initialement les articles "peer reviewed" dans HAL mais le processus était un peu fastidieux. Je pense que l'Université de Montpellier à laquelle est rattachée l'UMR MARBEC le fait maintenant pour les chercheurs.		On pourrait aussi envisager une petite sélection plus fine d'envoi de quelques textes : « ce mois-ci nos chercheurs ont publié ... ». Evidemment cela pose la question de la sélection des textes. Peut-être choisir un thème par mois ? Ou un département par mois ? Cela ferait circuler l'information scientifique (sans mettre les équipes en compétition bien sûr). Avec des portails comme sci-hub ou ResearchGate, même si je comprends que cela pose des questions de droits d'auteurs, j'aurais qu'il m'arrive en réalité rarement de télécharger des versions d'articles depuis HAL ou ARCHIMER.
Chercheur 46	Centralisation des papiers	pratique pour les évaluations biennales	Malheureusement non, je ne dépose pas dans HAL [procédures lourdes de tout redéposer les métadonnées]. Il y aurait un envoi automatique depuis Horizon cela ne me poserait pas problème. Les références sont également ajoutées systématiquement au site du Ceped (références par les PDR), qui prend aussi en compte les communications orales et poster. De plus, je maintiens une base Zotero pour mon site personnel	https://joseph.larmarange.net/spip.php?page=biblio (note : il s'agit du même outil Zotspip que celui utilisé pour le Ceped)	Merci!