

couperin.org

Consortium Unifié des Établissements Universitaires et
de Recherche pour l'Accès aux Publications Numériques

Shibboleth et modalités d'accès distants : une évaluation par les établissements du consortium

Enquête du consortium Couperin.org
Printemps 2018

Ce document est la restitution de l'enquête menée par le consortium Couperin.org entre le 26 avril et le 25 mai 2018 et pour laquelle 99 réponses ont été collectées.

Résumé Exécutif	3
Les questions et les réponses détaillées	4
Q1 - Qui êtes-vous ?	4
Q2 - Votre établissement donne actuellement accès aux ressources documentaires numériques via (plusieurs choix possible)	7
Q3 - L'implémentation du protocole Shibboleth par un fournisseur qui ne le propose pas encore serait :	8
Q4 - Dans le cas où vous estimez que l'implémentation du protocole Shibboleth est essentielle et prioritaire, pouvez-vous en détailler la ou les raisons ?	9
Q5 - Y'a-t-il d'autres fonctionnalités que vous souhaiteriez voir développer prioritairement chez les fournisseurs ? (par ex. : listes KBART, Rapports COUNTER 5, etc.)	11
Message d'invitation à l'enquête	12

Résumé Exécutif

Après avoir mené une première enquête en 2007¹ et organisé une journée d'étude en 2009², le consortium Couperin.org a souhaité ré-évaluer l'intérêt de ses membres pour la solution Shibboleth comme modalité d'accès aux ressources numériques pour pouvoir relayer auprès des éditeurs, qui nous en ont plusieurs fois fait la demande, l'état du besoin. Nous avons donc mené au printemps dernier une enquête au cours de laquelle **99 réponses**, pour **80 établissements distincts** (universités, grandes écoles, instituts de recherche, etc.) nous sont parvenues.

Sans grande surprise, **les solutions techniques continuent de coexister**. Si certains privilégient **Shibboleth**, en raison de la facilité d'accès procurée aux utilisateurs (lecteurs) et sa relative légèreté de mise en place, d'autres lui préfèrent l'utilisation d'un **reverse-proxy** qui permet aussi la collecte et l'analyse des traces de passage pour dresser une cartographie précise de l'usage des ressources numériques onéreuses³. Le choix de l'une ou l'autre des solutions dépend autant des préférences, disponibilité et connaissances des services informatiques que des besoins exprimés par les services documentaires. Il en résulte que **les modes d'accès sont en fait très souvent panachés au sein d'un même établissement**, en fonction des fournisseurs et des publics concernés. Ce qui complique d'autant la compréhension que peuvent avoir les usagers de ce qui leur permet l'accès ou non aux ressources qui les intéressent.

Pour donner accès en local (c'est-à-dire dans l'établissement), la **déclaration des plages IP** reste la méthode la plus utilisée, suivie par le reverse-proxy. Pour l'accès distant⁴, c'est désormais la solution d'un **reverse-proxy** qui domine, même si **Shibboleth** est aussi bien présent chez un tiers des répondants. La solution **VPN**, quoique minoritaire, reste utilisée.

Nombreux sont les collègues qui jugent que l'implémentation du protocole **Shibboleth** par un éditeur qui ne le propose pas déjà serait soit inutile soit non prioritaire, surtout par rapport à la **fourniture de statistiques d'usage conformes à COUNTER** ou des **listes de titres KBART** à jour et fournies par établissement, jugée largement prioritaire. L'absence d'informations qui aurait permis un monitoring fin des usages coté établissement⁵ est jugé rédhibitoire par beaucoup.

La sécurité et le confort procurés par **Shibboleth** à l'utilisateur sont toutefois indéniables, car ils permettent une navigation sans intermédiaire sur le site du fournisseur, indépendamment du lieu où l'utilisateur se situe. Le caractère prioritaire de l'implémentation de Shibboleth par un éditeur reste ainsi justifié pour une partie significative des répondants :

- stabilité des URLs, qui restent celles des éditeurs, à l'inverse d'un proxy qui nécessite une réécriture
- facilité de maintenance pour la DSI de l'établissement une fois la mise en place effectuée
- accès fluides et accélérés : il n'y pas de point de passage obligé par le reverse-proxy, qui peut parfois représenter un goulot d'étranglement

¹ https://www.couperin.org/IMG/pdf/Questionnaire_Acces_distant_Synthese.pdf

²

<https://www.couperin.org/services-et-prospective/acces-aux-ressources-cat/acces-distant/item/580-presentations-de-lapres-midi-couperin-afederation-didentites-et-acces-aux-ressources>

³ chantier dans lequel Couperin.org est très investi depuis la fin 2012, par le biais des outils ezPAARSE et ezMESURE

⁴ en plus de l'accès local ou non

⁵ car l'utilisateur une fois authentifié ne transite plus par un dispositif intermédiaire tel qu'un reverse-proxy

En complément de cette restitution, nous proposons aussi les éléments suivants :

1. Il est à noter aussi que l'installation et la maintenance d'un reverse-proxy représente une charge certaine pour les établissements qui s'en sont équipés puisqu'il faut prévoir d'appliquer régulièrement et rapidement des ajustements de paramètres⁶ ("stanzas") pour les différentes plateformes des fournisseurs.
2. A l'époque du questionnaire, LeanLibrary n'avait pas encore proposé d'offre pour son **extension de navigateur** en France à destination des membres du consortium. Cette extension, dont il existe aussi des versions libres et gratuites déjà mises en place dans certains établissements⁷, permet de remédier au passage forcé de l'utilisateur par les pages de l'établissement pour connaître les versions proxifiées des liens vers les ressources documentaires : en reconnaissant les noms de domaines concernés et en **redirigeant automatiquement vers le reverse-proxy**, l'extension rend le parcours proxifié relativement transparent pour l'utilisateur final qui l'installe et le configure dans son navigateur une fois pour toutes.
3. Nous rappelons enfin que le consortium Couperin, bien qu'engagé sur le terrain de la génération par les établissements de statistiques d'usage issues des traces de la proxification, encourage aussi les éditeurs à se conformer aux bonnes pratiques en vigueur, en particulier en ce qui concerne COUNTER :
 - une clause COUNTER est présente dans le contrat type. Les négociatrices et les négociateurs sont sensibilisés à cette question
 - nous sommes impliqués dans le travail des instances COUNTER (participation au comité exécutif et au comité technique)
 - nous avons fait traduire en français des documents de référence et d'accompagnement (guides conviviaux)
 - nous accompagnons les éditeurs qui le souhaitent (par le biais du Groupe de Travail sur les Indicateurs, GTI)
 - nous avons démarché divers opérateurs pour qu'une offre commerciale d'audit francophone indépendant puisse être proposée⁸

Les questions et les réponses détaillées

Q1 - Qui êtes-vous ?

Nous avons reçu **99 réponses**, pour **80 établissements distincts**, listés ci-dessous :

Institut Catholique de Paris	Université du Littoral Côte d'Opale
BIU Cujas	ENSAM
SCD Paris Sud	Université PSL

⁶ les échanges réguliers qui ont lieu sur la liste <https://groupes.renater.fr/sympa/info/code2bib> en sont le reflet

⁷ comme l'extension LAMA, proposée par l'Université d'Aix Marseille à ses publics : <https://github.com/SCD-Aix-Marseille-Universite/lama-addons>

⁸ une offre commerciale est proposée par Dekra et décrite sur <https://www.dekra-certification.fr/vos-attestations-et-ou-evaluations/l-audit-de-conformite-au-code-de-bonnes-pratiques-counter-dekra-certification.html>

Université de Caen Normandie	ENS Paris-Saclay
BIU de Montpellier (Université de Montpellier, Université Paul Valéry - Montpellier 3)	Campus Condorcet
SCD Le Mans université	Université d'Orléans
INHA (Institut national d'histoire de l'art)	SCD Paris Dauphine
Université Limoges	Université de Bourgogne (SCD)
EHESP	BULAC
Université Paris Nanterre	INSA Lyon
SCD université Paris 1	SCD de l'Université de Rennes 1
HEC	Université de Bretagne Occidentale
emlyon business school	Université d'Evry Val d'Essonne
Université de Corse	Université Toulouse III - Paul Sabatier
Genes Ensai	CNRS
Inria	Université Sorbonne nouvelle Paris 3 - DBU
IFREMER	Ecole Centrale de Lyon
Université de Cergy	IMT Atlantique
INSA Rouen Normandie	ENSTA Bretagne
Université Lille (ST)	Université de Pau et des Pays de l'Adour
Université de Reims	SCDU ANTILLES
SCD Paris 8	SCD Université de La Réunion
Université Claude Bernard Lyon 1	Université de Strasbourg (documentation électronique))
Université de Valenciennes et du Hainaut-Cambrésis	BSG
Université de Guyane	Université Clermont Auvergne
Université Jean Moulin Lyon3	École des Ponts ParisTech
Université de la Polynésie française	Université de TOURS
Bibliothèque interuniversitaire de la Sorbonne	Université Bretagne Sud
Bibliothèque Sainte-Geneviève	Université catholique de Lyon
Bibliothèques de l'ENS Paris	Université de technologie de Troyes

Collège de France	AUDENCIA BUSINESS SCHOOL
Grenoble Ecole de Management	Université de Lille
École nationale supérieure de mécanique et des microtechniques	SCD Université de Franche-Comté
Université Jean Monnet	Bu Paris 2
Université Lumière Lyon 2	Université de Reims Champagne-Ardenne
Université de la Nouvelle-Calédonie	Université Angers
Université Paris Descartes	IRD
SCD Université de Poitiers	Bibliothèque du CIRM
Mathdoc	Université de Nantes
ENS de Lyon - Bibliothèque Diderot de Lyon	

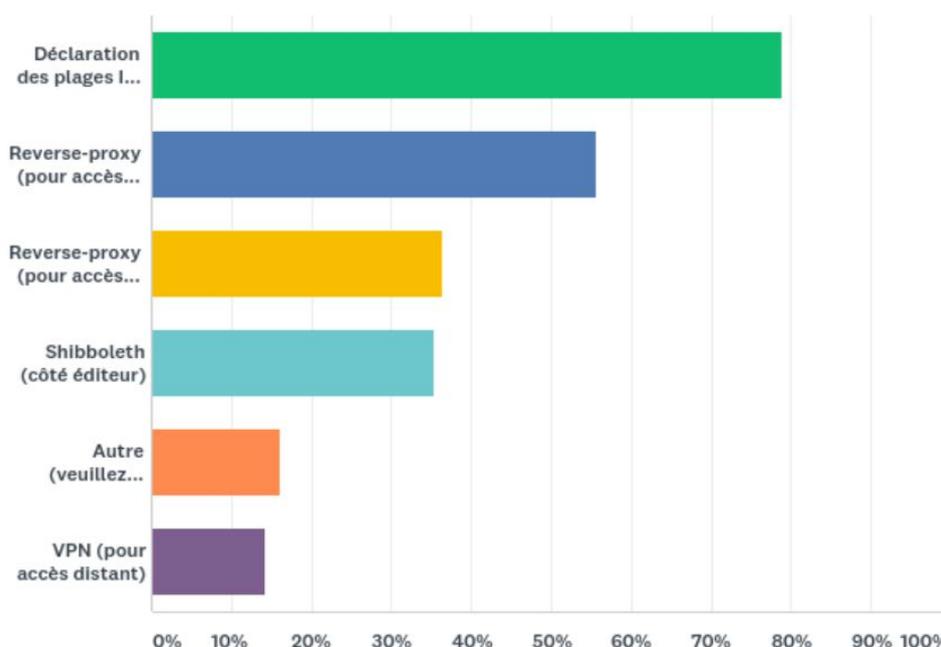
Q2 - Votre établissement donne actuellement accès aux ressources documentaires numériques via (plusieurs choix possible)

Plusieurs choix simultanés étaient possibles et les réponses sont réparties de la façon suivante :

- Déclaration des plages IP (pour accès local) → 78 réponses
- Reverse-proxy (pour accès local ET distant) → 55 réponses
- Reverse-proxy (pour accès distant seulement) → 36 réponses
- Shibboleth (côté éditeur) → 35 réponses
- VPN (pour accès distant) → 14 réponses
- Autre (veuillez préciser) → 16 réponses

Parmi les réponses "Autre", nous pouvons lister :

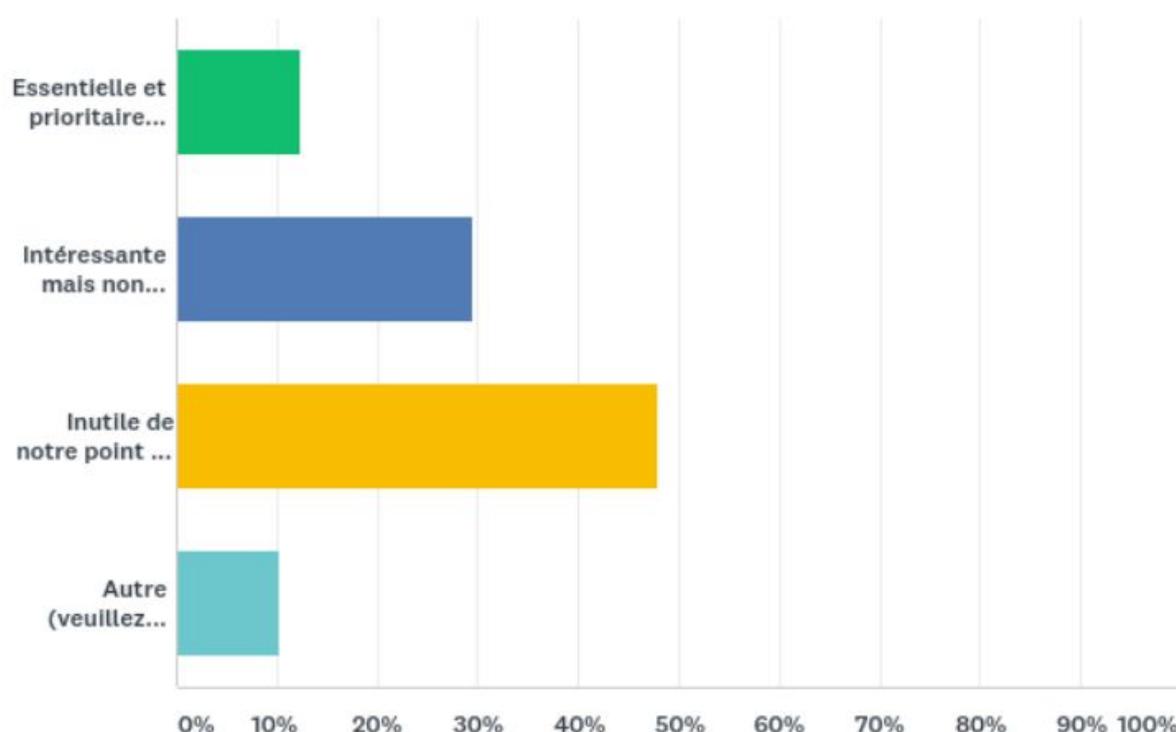
- SSO,
- Reverse-proxy + CAS,
- Reverse-proxy + Shibboleth
- Authentification interne Shibboleth pour les ressources produites par l'université,
- Mots de passe pour titres de revues isolés,
- VmWare,
- Proxy préconisé pour accès distant mais utilisable en local, accès local uniquement pour création de comptes auprès d'éditeurs particuliers,
- Certaines ressources en IP locales + reverse proxy, d'autres en reverse proxy seulement / nous ne faisons pas de pub pour Shibboleth mais savons que l'accès étant possible sur certaines ressources, certains usagers l'utilisent. Pour certaines ressources plus "artisanales" nous mettons à disposition identifiants et mots de passe dans l'ENT
- à titre exceptionnel : Shibboleth pour une ressource particulière (en droit) et certains enseignants utilisent un VPN



Remarque complémentaire : certains établissements nous ont indiqué privilégier l'accès via le reverse-proxy, tout en demandant également aux fournisseurs de paramétrer l'accès Shibboleth quand ils le proposent, mais uniquement en "roue de secours".

Q3 - L'implémentation du protocole Shibboleth par un fournisseur qui ne le propose pas encore serait :

- Essentielle et prioritaire (malgré les restrictions évoquées plus haut à propos de la collecte de traces de passage en local rendue impossible) - (12 réponses)
- Intéressante mais non prioritaire (29 réponses)
- Inutile de notre point de vue (47 réponses)
- Autre (veuillez préciser) (10 réponses)



Avant de procéder à l'analyse des réponses pour cette question, il nous faut signaler que certains d'entre vous nous ont reproché le caractère biaisé de cette question. Notre objectif n'était ni d'orienter la réponse ni de disqualifier la solution Shibboleth mais plutôt de rappeler la restriction (dont l'explication introductive aurait peut-être été hâtivement lue par certain.e.s répondant.e.s). Nous reconnaissons sans mal que cette intention était maladroitement formulée et nous vous prions de nous en excuser. Nous tâchons de tenir compte de cet élément dans cette restitution.

Parmi les commentaires que vous avez laissés figurent les éléments suivants :

- Les statistiques du reverse-proxy sont extrêmement précieuses. Le protocole Shibboleth peut être intéressant (car il ne nécessite pas de maintenance) mais ne devrait jamais suppléer l'accès par reverse-proxy, car dans ce cas de figure nous perdons toutes les statistiques d'usage provenant de nos propres sources et disposons uniquement des statistiques fournies par l'éditeur - ce qui rend quasiment impossible de comparer les éditeurs.

- Shibboleth peut permettre de recréer ce qu'on obtient avec le reverse proxy en envoyant des données utilisateurs anonymisées aux fournisseurs. En utilisant cette potentialité, le confort pour l'utilisateur lié à la connexion directement depuis la plateforme est très importante. Donc, l'implémentation du protocole pourrait être capitale mais dans sa configuration optimale
- Ne sais pas, et le service informatique ne connaît rien sur le sujet
- Inutile et dangereux (partage de données utilisateurs avec un tiers). Nos RSSI et CIL ne seraient vraisemblablement pas d'accord.
- Mal venu : si on fait du Shibboleth, on perd les statistiques car on perd la collecte des traces.
- Intéressant sous réserve de ressources humaines suffisantes à la Direction Informatique
- Serait intéressant si TOUS les éditeurs le faisaient car on pourrait faire l'économie de licences EZproxy; MAIS on perdrait les apports actuels d'EZparse, qui sont essentiels pour estimer correctement les usages actuels de notre doc élec. Sans cette étape reverse proxy, on dépendrait à nouveau entièrement des statistiques fournies par les éditeurs (sauf nouvel outil inconnu de nos services actuellement)
- Question biaisée à partir du moment où les établissements sont encouragés depuis plusieurs années à mettre en place EZParse, et donc à privilégier le reverse-proxy pour l'accès local ET distant. Nous sommes en cours d'implantation d'EZParse, et donc nous ne favorisons plus Shibboleth, pour cette seule raison.
- Cela dépend des fournisseurs
- Toujours intéressant pour les usagers de faciliter l'accès aux ressources par divers moyens, mais préjudiciables pour un monitoring fin des logs de notre côté

Q4 - Dans le cas où vous estimez que l'implémentation du protocole Shibboleth est essentielle et prioritaire, pouvez-vous en détailler la ou les raisons ?

Nous avons reçu ici une petite vingtaine de réponses (soit un peu plus que les 12 qui avaient répondu que l'implémentation du protocole Shibboleth était essentielle et prioritaire à la question précédente) :

- Notre DSI a longtemps espéré contourner la solution reverse-proxy par l'utilisation de Shibboleth ou l'installation d'un http referer. La raison principale invoquée (outre les problèmes de définition de priorités propres à la DSI) était de rendre plus fluide et rapide la consultation des ressources électroniques. Vu le manque d'homogénéité sur les modalités d'accès distant selon les éditeurs, et les difficultés de communiquer de façon claire sur ces

modalités différentes auprès de nos usagers, nous avons finalement opté il y a 1 an pour la solution EZ Proxy.

- Si demain tous les éditeurs proposaient le protocole Shibboleth, nous nous poserions tout de même la question de l'abandon du reverse-proxy. J'y vois un autre avantage, celui de nous éviter les mises à jours et différents paramétrages à effectuer dans le reverse proxy, surtout lorsque au sein de la bibliothèque, nous n'avons pas la main sur l'outil et qu'il faut passer par une tierce personne. Toutefois, il faudrait assurément pouvoir garder une trace du passage de nos utilisateurs sur la ressource.
- Réponse DSI: nous avons déjà mis en place cette technologie
- Pas d'avis pour reverse proxy ou shibboleth. La seule chose que je sais : actuellement le VPN ne permet l'accès qu'à trop peu de ressources, et c'est hétérogène d'un usager à l'autre. Catastrophique. Pas de résolution côté service informatique
- gestion simplifiée pour les accès et aussi pour les gestionnaires côté DSI
- Nous avons implémenté depuis 3 ans l'accès à certaines ressources numériques avec Shibboleth (celles qui étaient compatibles). Nous ne disposions jusqu'alors que d'accès par plage d'adresses IP (sans reverse proxy) ou par création de comptes manuels pour des usages spécifiques. C'est pourquoi nous privilégions Shibboleth.
- Nous ne passons que par EZproxy, qui lui passe par Shibboleth. Cela nous permet d'avoir des logs enrichis.
- Principalement une raison d'utilisabilité et de "transparence" du portillon. Le fait de collecter des données sur les usages ne doit pas se faire aux dépens du confort des usagers. En revanche, ce qui complique la vie des gens, c'est la cohabitation de plusieurs systèmes de connexion très différents. Le proxy permet d'unifier l'expérience utilisateur de ce point de vue, raison pour laquelle nous avons fait ce choix technique en 2009.
- C'est une façon efficace d'avoir un accès à distance quand on a l'habitude d'utiliser l'URL de l'éditeur. Il n'est pas nécessaire de donner des URL spécifiques. Le VPN n'est pas satisfaisant pour nous.
- Note: votre question 3 est biaisée. Le "malgré les restrictions..." influence la réponse. Chaque solution a ses désavantages. La raison principale est le service rendu aux lecteurs, qui est meilleur avec une authentification unique.
Déclaration IP: uniquement au sein d'un laboratoire.
Reverse-proxy: gros investissement pour que cela soit fait correctement. Dans notre cas, il y a des licences locales (signées par des bibliothèques d'unités), des licences d'universités (gérées par les SCD), des licences nationales (gérées par le RNBM, l'INIST, l'ABES). Un seul reverse-proxy ne peut pas tout gérer. Dans certaines universités, le reverse proxy du SCD ne gère pas les licences locales et refuse que l'IP d'un reverse-proxy local soit déclarée pour les licences nationales. Résultat: le chercheur doit se connecter sur plusieurs sites pour accéder aux revues et il doit savoir quel site lui permet d'accéder à quels articles (dans de nombreux cas, les archives sont à un endroit, le courant à un autre endroit).
Pour que la navigation fonctionne de bout en bout, il faut déclarer dans la configuration du reverse-proxy les URLs de toutes les ressources documentaires en accès libres (+ Google Scholar...). Sinon les URLs ne sont pas ré-écrites lors des rebonds et on va vite sortir du chemin avec le reverse-proxy en intermédiaire. Il faut alors une extension de navigateur pour revenir dans le bon chemin.
Avec une authentification unique, le lecteur n'a plus à se préoccuper de comment accéder à la ressource. Il a une citation, un DOI, il arrive sur le site de la revue et ça marche (ou pas s'il n'a pas les droits).
Les stats d'usage DEVRAIENT être négligeables en comparaison au service rendu aux

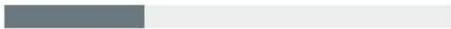
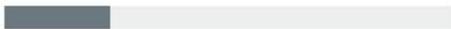
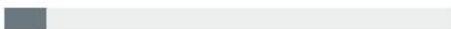
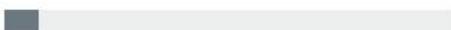
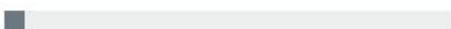
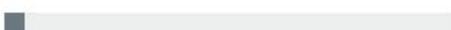
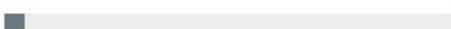
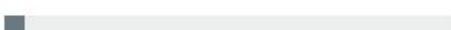
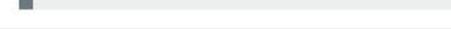
lecteurs.

- Nous disposons, dans l'immédiat, d'un IdP Shibboleth opérationnel. Une tentative de reverse-proxy s'est avérée "non satisfaisante" pour nos usages et n'est donc plus implémentée. VPN n'est pas une solution adaptée à notre cas.
- Shibboleth propose à la fois la solution la plus commode pour l'utilisateur et la plus conforme aux règles de sécurité internet, en particulier comparé à une solution de reverse proxy.
- Shibboleth simplifie et sécurise la gestion des accès indépendamment du lieu où le client se situe.
- la DSI de l'université ne jure que par Shibboleth. aucune discussion possible et blocage de l'accès distant actuellement pour certaines ressources qui ne sont pas compatibles avec Shibboleth...
- Shibboleth est une demande du service informatique de l'établissement, qui s'inscrit dans le cadre de RENATER et, pour la solution reverse-proxy, peine souvent à intégrer rapidement les changements nécessaires

Q5 - Y'a-t-il d'autres fonctionnalités que vous souhaiteriez voir développer prioritairement chez les fournisseurs ? (par ex. : listes KBART, Rapports COUNTER 5, etc.)

On voit ici que la fourniture des listes **Kbart** (ou Fichiers Kbart) et des **rapports COUNTER** (rapports statistiques, statistiques d'usage) par les éditeurs sont les souhaits de fonctionnalités les plus souvent exprimés.

Sont mentionnés, plus ponctuellement toutefois : API, SUDOC, outils de découverte

Listes KBART		32,81%	21
Rapports Counter 5		31,25%	20
Éditeurs		23,44%	15
Fourniture		9,38%	6
Fichiers KBART		7,81%	5
Rapports Statistiques		4,69%	3
Données		4,69%	3
D'usage		4,69%	3
Téléchargeables		4,69%	3
Outils de Découverte		3,13%	2
API		3,13%	2
SUDOC		3,13%	2

Message d'invitation à l'enquête

Envoyé aux listes :

correspondants_services_associes@listes.couperin.org

achats@listes.couperin.org

code2bib@groupe.renater.fr

ezpaarse-users@groupe.renater.fr

Bonjour à tou.te.s,

Le consortium a récemment été contacté par l'équipe technique d'un éditeur au sujet d'une éventuelle implémentation de Shibboleth sur leur plate-forme. Ils en sont actuellement à la fin de l'analyse de faisabilité.

Nous avons émis la réserve suivante : la collecte des traces de passage, nécessaire pour l'utilisation d'ezPAARSE par exemple, devient impossible quand un éditeur se déclare comme "Service Provider" auprès de la fédération d'identités. L'utilisateur, une fois authentifié, n'est plus en interaction qu'avec le serveur de l'éditeur et la trace de son activité de consultation sur cette plateforme est perdue.

Côté Couperin, nous souhaiterions à cette occasion ré-évaluer l'intérêt de Shibboleth comme modalité d'accès aux ressources numériques. Nous croyons savoir que certains établissements la privilégient, en raison de sa facilité d'accès pour les utilisateurs et sa relative légèreté de mise en place ; tandis que d'autres lui préfèrent l'utilisation d'un reverse-proxy, cas de figure le plus commun actuellement.

Pour rappel des éléments techniques, nous vous invitons à consulter la page dédiée du site Couperin.org :

<https://www.couperin.org/services-et-prospective/acces-aux-ressources-cat/acces-distant/item/644-lacces-distant-shibboleth-et-les-federations-didentites>

concernant les modalités d'accès distant, et en particulier la section qui concerne les statistiques.

L'enquête se trouve sur :

https://fr.surveymonkey.com/r/shibboleth_couperin

Elle sera clôturée à la fin du mois de mai.