

Comprendre l'usage d'un outil de découverte, par les données : l'exemple de swisscovery HES-SO

Thomas KERBOUL
thomas.kerboul@geneve.ch
<https://orcid.org/0009-0005-0186-3267>
Spécialiste métadonnées, Bibliothèque de Genève

Delphine LEBOULEUR
delphine.leboulleur@gmail.com
<https://orcid.org/0009-0003-3940-6227>
Spécialiste en Sciences de l'information

Résumé

Trois ans après l'implémentation de swisscovery, notre recherche vise à comprendre l'utilisation de cet outil de découverte par la communauté académique de la Haute école spécialisée de Suisse occidentale (HES-SO). Elle s'appuie sur les statistiques d'usage issues des modules Analytics de Primo VE et d'Alma, logiciels d'Ex Libris constituant swisscovery, ainsi que sur les données COUNTER de deux fournisseurs de ressources électroniques, produites sur 2022. Cette étude descriptive et exploratoire a notamment permis de caractériser un parcours utilisateur moyen, visualisé sous la forme d'un diagramme de Sankey, dont les indicateurs nécessaires à sa construction sont présentés ici. Malgré certaines limites d'exploitabilité des données, notre travail conclut à un usage contrasté de swisscovery, souvent limité à des recherches documentaires basiques, et soumis à la très forte concurrence d'autres moyens d'accès aux ressources électroniques.

Mots-clés

outils de découverte, Primo, bibliothèques académiques, données d'usage, indicateurs, pratiques informationnelles

Abstract

Three years after the implementation of swisscovery, our research aims to understand the usage of this discovery tool by the academic community of the University of applied sciences and arts western Switzerland (HES-SO). It relies on usage statistics from the Analytics modules of Primo VE and Alma, the Ex Libris software components of swisscovery, as well as COUNTER data from two electronic resource providers, generated in 2022. This descriptive and exploratory study has allowed us to characterize an average user journey, visualized in the form of a Sankey diagram, with the necessary indicators for its construction presented here. Despite certain limitations in data exploitability, our work concludes with a contrasting usage pattern, often limited to basic documentary searches, and subject to strong competition from other access tools to electronic resources.

Keywords

discovery tools, Primo, academic libraries, usage data, indicator, information behavior

Zusammenfassung

Drei Jahre nach der Implementierung von swisscovery zielt unsere Forschung darauf ab, die Nutzung dieses Entdeckungstools durch die akademische Gemeinschaft der Fachhochschule Westschweiz (HES-SO) zu verstehen. Sie stützt sich auf Nutzungsstatistiken der Analytics-Module von Primo VE und Alma, den Ex Libris-Softwarekomponenten von swisscovery, sowie auf COUNTER-Daten von zwei Anbietern elektronischer Ressourcen, die im Jahr 2022 generiert wurden. Diese beschreibende und explorative Studie hat es uns ermöglicht, einen durchschnittlichen Nutzerweg zu charakterisieren, der in Form eines Sankey-Diagramms visualisiert wird, wobei die für seine Erstellung erforderlichen Indikatoren hier präsentiert werden. Trotz gewisser Einschränkungen bei der Datenverwertung kommt unsere Arbeit zu einem kontrastreichen Nutzungsmuster, das oft auf grundlegende dokumentarische Suchen beschränkt ist und einer starken Konkurrenz durch andere Zugriffsmöglichkeiten auf elektronische Ressourcen unterliegt.

Schlagwort

Discovery Service, Primo, Universitätsbibliotheken, Nutzungsdaten, Indikator, Informationsverhalten

Abstract

Tre anni dopo l'implementazione di swisscovery, la nostra ricerca mira a comprendere l'utilizzo di questo strumento di scoperta da parte della comunità accademica della Scuola universitaria professionale della Svizzera occidentale (HES-SO). Si basa su statistiche di utilizzo provenienti dai moduli Analytics di Primo VE e Alma, i software di Ex Libris che costituiscono swisscovery, oltre che sui dati COUNTER di due fornitori di risorse elettroniche, prodotti nel 2022. Questo studio descrittivo ed esplorativo ci ha permesso di caratterizzare un percorso medio dell'utente, visualizzato sotto forma di un diagramma di Sankey, con gli indicatori necessari per la sua costruzione presentati qui. Nonostante alcune limitazioni nell'exploitabilità dei dati, il nostro lavoro conclude con un modello d'uso contrastante, spesso limitato a ricerche documentarie di base e soggetto a una forte concorrenza da parte di altri strumenti di accesso alle risorse elettroniche.

Descrittore

discovery service, Primo, biblioteche universitarie, dati di utilizzo, indicatore, comportamento informativo



Cet article est disponible sous licence [Creative Commons Attribution - Partage dans les Mêmes Conditions 4.0 International](https://creativecommons.org/licenses/by-sa/4.0/).

1. Introduction

Swisscovery est une plateforme fédérant les catalogues de 500 bibliothèques universitaires et spécialisées de Suisse. Inauguré en décembre 2020 par SLSP (Swiss Library Service Platform), cet outil de découverte, composé des logiciels Alma et Primo VE d'Ex Libris, a ainsi remplacé six réseaux de bibliothèques existants et offre désormais l'accès à près de 41,6 millions de documents (SLSP, 2023).

Dans le cadre d'un projet de recherche réalisé à la Haute école de gestion de Genève pour le master en Sciences de l'Information et sous le mandat de Centre de l'Information Scientifique HES-SO (CISO), nous nous sommes intéressés à l'usage qui était fait de swisscovery par la communauté académique de la Haute école spécialisée de Suisse occidentale (HES-SO), un réseau de vingt-huit hautes écoles situées en Suisse romande.

Notre étude avait pour objectif d'établir un bilan de l'usage de swisscovery HES-SO, puisque, à notre connaissance et maintenant trois ans après son lancement, aucun rapport national chiffré sur l'utilisation ou la performance de swisscovery n'a encore été rendu public. Malgré les plus de 24 millions de connexions en Suisse rapportés par SLSP en 2022 (Heinemann et al., 2023, p. 97), il nous apparaissait essentiel de mieux comprendre la place que cet outil occupe dans les pratiques informationnelles de ses utilisateurs-trices potentiel-les, entendues comme les moyens mis en œuvre pour accéder à l'information dont ils ou elles ont besoin.

Dans ce contexte, nous avons choisi de présenter ici un des livrables majeurs de notre projet, à savoir la construction d'un diagramme de Sankey du parcours moyen des utilisateurs-trices de la plateforme en 2022. Il présente l'intérêt d'apporter, de façon synthétique, un éclairage sur l'usage de swisscovery dans les pratiques informationnelles de la communauté académique de la HES-SO.

Cet article vise ainsi à partager notre démarche, dans le but de faciliter l'exploration des statistiques d'usage par d'autres institutions du réseau swisscovery ; de proposer des indicateurs simples d'évaluation des pratiques informationnelles de leur population cible ; et enfin d'initier une discussion sur leur pertinence et sur les résultats obtenus, sous le prisme d'autres études académiques. En revanche, nous ne formulons pas d'hypothèses sur les niveaux de performance attendus ni sur l'évolution des indicateurs produits.

2. Revue de la littérature

Des études menées dans d'autres pays ont mis en lumière l'impact contrasté que le passage à un outil de découverte avait pu générer sur l'utilisation des ressources électroniques (Ngo et al., 2019). En effet, les outils de découverte, qui ont remplacé les catalogues traditionnels des bibliothèques, offrent l'avantage d'élargir la portée de la recherche documentaire non seulement aux collections physiques des bibliothèques, mais aussi à leurs collections numériques, aux bases de données d'articles issues de divers fournisseurs, aux pages web et aux dépôts institutionnels. Ils permettent ainsi de faire face à l'explosion du nombre de ressources électroniques et de proposer une recherche exploratoire facilitée par des fonctionnalités de navigation proches de celles des moteurs de recherche (Breeding, 2009; Jiang et al., 2017).

Néanmoins, accessibles via un navigateur internet, ces outils se retrouvent en concurrence directe avec d'autres moteurs de recherche de données bibliographiques, tels que Google

Scholar, ou des bases de données spécialisées comme MEDLINE via PubMed ou JSTOR (Wu & Chen, 2014; Namei & Young, 2015; Nicholas et al., 2017; Cummings, 2021; Alotaibi et al., 2022; et en particulier pour swisscovery Merminod, 2023, p. 10). Ce phénomène est encore amplifié par la généralisation de l'Open Access, promue en Suisse par swissuniversities, qui vise 100% de publications en libre accès à horizon 2024 (Werlen, 2017) et interroge d'autant plus les modalités d'accès à ces ressources ouvertes qu'elles peuvent s'avérer plus difficilement accessibles via les outils de découverte (Bullock, 2021).

Parmi la littérature qui s'intéresse à l'utilisation des outils de découverte, on peut relever deux courants principaux. Le premier s'intéresse à l'outil de découverte en tant que tel et au comportement des utilisateurs-trices sur celui-ci, en compilant différentes statistiques ; au contraire d'études basées sur des questionnaires ou des entretiens, l'approche est donc essentiellement quantitative.

Ainsi, Durante et Wang (2012) ont utilisé les données de Google Analytics pour élaborer un cadre d'évaluation de leur outil de découverte (Primo) reposant sur quatre catégories d'indicateurs : le nombre d'utilisateurs-trices, leur taux de satisfaction, leur engagement avec les fonctionnalités de la plateforme et le taux de réussite de leurs recherches.

Huttenlock et Malone (2013) se sont quant à eux basés sur l'analyse des fichiers journaux de l'outil de découverte (WorldCat Local) pour calculer plusieurs métriques d'intérêt : le point de départ d'une recherche, le temps passé sur une page, le temps d'une visite, l'action suivant une recherche ou l'affichage d'un résultat, l'usage des facettes, le nombre de résultats par recherche ainsi que le taux de succès d'une recherche, entendu à la fois comme le nombre de requêtes ayant affiché des résultats et la sélection d'un résultat après une recherche, ainsi que les stratégies de recherche sur l'outil de découverte. L'utilisation de Google Analytics leur permet en outre de trianguler leurs résultats.

De leur côté, Cohen et Thorpe (2015) se sont intéressées aux données générées par l'outil de découverte (EBSCO Discovery Service) pour connaître le nombre de recherches, le nombre de sessions, le nombre de visualisation des résumés et d'accès au texte intégral depuis l'outil de découverte ainsi que via le résolveur de liens. Elles ont également eu recours à Google Analytics pour calculer le taux d'utilisateurs-trices connus, c'est-à-dire ayant déjà visité le site par le passé, le type d'appareil et de navigateur utilisés ainsi que le texte des requêtes et les champs d'interrogation utilisés.

Moore (2016) a adopté une approche originale, basée sur les données fournies par l'outil de découverte (Primo), pour classer les actions des utilisateurs-trices en quatre catégories répondant à différentes intentions : obtenir la ressource de façon immédiate, souhaiter la consulter ultérieurement, être intéressé-e par plus de détails sur la ressource ou rechercher une ressource similaire.

Une étude qui nous a beaucoup inspiré est celle de Jiang et al. (2017), qui analysent les fichiers journaux pour réaliser un diagramme visuel représentant les chemins suivis par les utilisateurs-trices sur l'outil de découverte (Primo) entre cinq types de pages : la page d'accueil, la page de recherche, la page de résultats, la page de détail d'un document et la page personnelle de l'utilisateur-trice. Ils se sont également intéressés aux actions réalisées sur ces pages et aux caractéristiques des sessions enregistrées.

Enfin, spécialement sur swisscovery, Merminod (2023) a triangulé des tests utilisateur et des données d'observations en bibliothèque avec les données d'usage de l'outil de découverte (Primo VE). Elle s'est en particulier intéressée aux types de recherches, aux actions sur la page de résultats, y compris l'utilisation des facettes et du tri, ainsi qu'aux actions sur les notices.

Toutefois, les six études présentées ci-dessus se focalisent sur l'usage de l'outil de découverte et ne le lient pas à l'usage des collections ; elles ne permettent donc pas d'appréhender la place des outils de découverte dans les pratiques informationnelles des usagers-ères, prises dans leur globalité. C'est précisément à cette problématique que s'intéresse un second courant dans la littérature secondaire.

Dans deux études, les bibliothécaires de la J. Paul Leonard Library à San Francisco ont évalué l'utilisation des outils de recherche en examinant les données du résolveur de liens et du prêt interbibliothèques, à l'aide d'un proxy, pour déterminer si les utilisateurs-trices accèdent au contenu depuis la base de données de la bibliothèque, l'outil de découverte (OneSearch) ou Google Scholar (Wang & Howard, 2012; Wang, 2020).

Jouneau et Porquet (2019), affiliés au Consortium Couperin, adoptent une approche similaire mais basée exclusivement sur l'analyse des fichiers journaux grâce au logiciel ezPAARSE. Leur étude a toutefois nécessité que l'ensemble des utilisateurs-trices des services de la bibliothèque soit authentifié pour pouvoir évaluer leur utilisation de manière fine.

D'autres études se basent quant à elles sur les HTTP Referers fournis par les fournisseurs de contenu et analysent d'où provient le trafic ayant mené à un document (Evans & Schonfeld, 2020, p. 15; Cummings, 2021). Les HTTP Referers permettent en effet d'identifier d'où provenait l'utilisateur-trice avant d'atteindre le document qu'il désirait ; néanmoins, il s'agit de données qui ne sont pas habituellement communiquées par les fournisseurs de contenu.

Ces cinq études n'ont pas fait appel aux données de consultation traditionnelles des fournisseurs de contenu, nommément les rapports COUNTER, auxquels les bibliothécaires sont plus habitués. Jouneau et Porquet (2019) en particulier sont critiques sur ces données, selon eux trop agrégées et implémentées de manière trop hétérogène par les fournisseurs. On note notamment que ces études ont été réalisées avec un fort soutien informatique, permettant une collecte à grande échelle des actions effectuées par les utilisateurs-trices, ce qui n'est pas toujours applicable. Dans le cadre de notre projet de recherche, nous avons également demandé l'accès aux fichiers journaux, mais cette demande n'a pas abouti. Sur l'usage de ce type de données, Gustafson-Sundell (2023) indique des préoccupations liées à la confidentialité des données.

Cependant, une étude de S. Robins (2022) combine à la fois les rapports COUNTER et les données de téléchargement direct à travers EBSCO pour évaluer la part d'utilisateurs-trices passant par l'outil de découverte (EBSCO Discovery Service) pour atteindre les documents. S. Robins arrive à des conclusions qu'il juge lui-même surprenantes, car à contre-courant de la majorité de la littérature scientifique. En effet, il démontre qu'EBSCO est la plateforme préférée des utilisateurs-trices pour accéder au texte intégral. Toutefois, il souligne que la fonctionnalité d'accès au texte intégral directement depuis l'outil de découverte est certainement un facteur déterminant.

Bien que nous nous inscrivions pleinement dans la continuité des études présentées et particulièrement celle de S. Robins, l'originalité de notre étude a été de combiner les deux approches décrites précédemment, pour comprendre à la fois comment et à quel point l'outil de découverte est utilisé pour accéder aux ressources documentaires. Ainsi, outre l'utilisation de la plateforme en tant que telle, nous avons cherché à replacer son usage dans des pratiques informationnelles plus globales, et ceci au moyen de données standards auxquelles les bibliothécaires ont facilement accès.

3. Méthodologie et données

3.1. Démarche

Le cœur de notre étude a consisté en l'analyse de plusieurs corpus de données.

Dans un premier temps, nous nous sommes placés dans la continuité de Merminod (2023) et, sur conseil du Service de coordination de la Division de l'information scientifique (CODIS) de l'Université de Genève (UNIGE), nous avons collecté les données issues des rapports standards de Primo VE Analytics, dits « out-of-the-box ». En nous référant à la documentation fournie par Ex Libris (Ex Libris, 2023a), nous avons ensuite exploré ces données d'usage en formalisant une cartographie, afin d'ordonner les actions les unes par rapports aux autres. Ce corpus ne fournissant aucune donnée concernant l'accès aux documents, qu'ils soient physiques ou électroniques, nous l'avons complété par de nouvelles données issues d'Alma Analytics, concernant les prêts physiques, internes et externes, et le résolveur de liens. Enfin, afin de replacer l'usage de swisscovery dans l'ensemble des parcours informationnels possibles, l'usage global des ressources électroniques a été quantifié, en se limitant, pour des questions de faisabilité, à l'étude des données COUNTER de deux fournisseurs majeurs à la HES-SO, que sont Elsevier et Springer, et aux journaux en particulier.

Conjointement à l'analyse des données, nous avons mené une revue de littérature sur la production d'indicateurs de suivi de l'usage des outils de découverte. Outre les études mentionnées précédemment, nous nous sommes intéressés aux retours d'expérience des bibliothèques qui avaient mis en place un outil de découverte et en avaient mesuré l'impact. Nous avons également complété cette revue de littérature par l'étude des normes ISO 2789 (2022), ISO 11620 (2023) et ISO 16439 (2014).

En faisant la synthèse des indicateurs jugés pertinents par les publications consultées, et en excluant ceux nécessitant des données auxquelles nous n'avions pas accès, nous avons retenu une liste d'indicateurs, que nous avons calculés et formalisés dans des fiches descriptives respectant le format standard de présentation de la norme ISO 11620 (2023, p. 22-24). Nous y avons consigné l'ensemble des informations nécessaires à leur production, leur bonne interprétation et la prise en compte de leurs limites. Ces fiches descriptives, classés en cinq catégories, sont disponibles dans notre mémoire (Kerboul & Lebouleur, 2024a, p. 139-169).

En parallèle, nous avons recherché une représentation visuelle impactante pour nos résultats. L'étude de Jiang et al. (2017) proposait une visualisation sous la forme d'un diagramme de Sankey qui rendait immédiatement compréhensibles les parcours empruntés par les utilisateurs-trices. Nous nous en sommes inspirés pour représenter un parcours utilisateur type sur swisscovery, simplifié et macroscopique du fait des données annuelles agrégées

disponibles, que nous avons mis en perspective des ressources électroniques consultées via d'autres plateformes.

Nous avons triangulé nos résultats à chaque fois que cela a été possible avec les données de l'Office Fédéral de la Statistique (Office fédéral de la statistique (OFS), 2023: variables 2 et 21), ainsi qu'avec les résultats des études relevées dans la revue de littérature. Enfin, nous avons collaboré avec les membres du CISO, qui ont notamment extrait les rapports identifiés et relayé nos questions auprès de SLSP et d'Ex Libris lorsque cela s'est avéré nécessaire.

3.2. Population cible

Vingt-huit bibliothèques desservent la communauté de la HES-SO, affiliées et réparties de façon non uniforme sur les différents sites. Elles sont coordonnées par le CISO. Parmi ces bibliothèques, dix-sept sont membres de swisscovery (HES-SO, 2023), tandis que quatre sont membres de Renouvaud (HES-SO, 2022), le réseau de bibliothèques du canton de Vaud, et que les autres disposent d'un OPAC individuel.

Sur cette base, notre population cible a été estimée à 21'425 utilisateurs-trices potentiel-les de swisscovery à la HES-SO sur l'année 2022, soit 18'719 étudiant-e-s et 2'706 ETP (comprenant les enseignant.es avec responsabilités de direction pour une unité organisationnelle, les autres enseignant.es, les collaborateurs-trices scientifiques ainsi que les assistant-e-s et/ou doctorant-e-s), affilié-e-s aux hautes écoles dont les bibliothèques sont membres de swisscovery (Kerboul & Leboulleur, 2024a, p.23 et p.128).

3.3. Corpus de données

L'ensemble des rapports présentés ci-dessous a été extrait sur le périmètre de la HES-SO et pour l'année 2022. Pour l'essentiel, il s'agit de fichiers Excel rapportant différentes métriques généralement agrégées par mois. Dans les tableaux suivants [Tableaux 1, 2 et 3], les rapports grisés n'ont pas été exploités dans le cadre de notre recherche, les autres rapports ont tous été analysés dans notre mémoire et les rapports avec astérisque ont été nécessaires à la production du parcours utilisateur présenté dans cet article [Chapitre 4.2]. Nous renvoyons à notre mémoire pour une analyse détaillée du contenu de chaque rapport (Kerboul & Leboulleur, 2024b, p. 23-50).

3.3.1. Données Primo VE Analytics

Concernant les données Primo VE, nous avons eu accès à douze rapports standards, dits « out-of-the-box », couvrant six domaines différents [Tableau 1]. Ce dataset a été déposé en soutien à notre mémoire de recherche sur Zenodo (Kerboul & Leboulleur, 2024b).

Tableau 1 : Rapports standards « out-of-the-box » de Primo VE Analytics sur les données d'usage.

Rapports standards Primo VE	Subject areas
1. Actions per Session	Action Usage
2. Action Usage General	
3. Action Usage Searches *	
4. Action Usage Results List	
5. Action Usage Document *	
6. Action Usage Exploration *	
7. Action Usage	

8. Device Usage	Device Usage
9. Facet Usage	Facet Usage
10. Popular Searches	Popular Searches
11. Session *	Sessions
12. Zero Result Searches *	Zero Result Searches

Les rapports liés à la table *Action Usage* rendent compte des interactions des utilisatrices avec l'interface de swisscovery HES-SO, et en comptabilisant le nombre de clics pour chacune d'entre elles (*Actions*). Ces actions sont regroupées par groupes et sous-groupes, et génèrent des rapports *Action Usage* thématiques, en lien avec le processus de recherche documentaire : *General*, *Searches*, *Results List*, *Document* et *Exploration*. Nous n'avons pas utilisé le rapport *Action Usage*, qui ne fait que reprendre certaines données de chacun des rapports précités, ni le rapport *Actions per Session*, auquel nous avons préféré le rapport *Session* pour des raisons que nous évoquerons plus loin.

Le rapport *Zero Result Searches* fournit quant à lui des informations qualitatives et quantitatives sur les recherches, basiques et avancées, qui n'ont pas retourné de résultat.

Enfin le rapport *Session* comptabilise le nombre de sessions, d'actions, d'identifications, ainsi que la durée moyenne des sessions, par pays de localisation de l'adresse IP de l'utilisatrice.

3.3.2. Données Alma Analytics

Les données complémentaires collectées d'Alma Analytics sont issues des rapports suivants [Tableau 2].

Tableau 1 : Rapports complémentaires issus d'Alma Analytics.

Rapports Alma
1. Active User per loan (per library)
2. Count of Items Loaned by Patron Group *
3. Ressource Sharing Borrowing Requests_Per Partner
4. Resolver Statistics by Source *

Concernant les documents physiques, le rapport *Count of Items Loaned by Patron Group* comptabilise les prêts internes (*In House*) et externes (*Not In House*), dans les bibliothèques de la HES-SO membres du réseau swisscovery. Le prêt interne sert à tracer les documents qui sont utilisés dans la bibliothèque mais qui ne sont pas empruntés.

Pour les ressources électroniques, nous avons utilisé le rapport sur le résolveur de liens *Resolver Statistics by Source*, initialement créé par Ex Libris et augmenté de quelques colonnes pour répondre à notre besoin (*Source*, *Material Type*, *Source Type*, *Normalized Source*, *Number of Requests*, *Number of Clicked Requests*, *Number of Requests Without Services* et *Electronic Collection ID*). Le résolveur de liens établit un lien direct entre les différents outils permettant de localiser les ressources électroniques (swisscovery, Google Scholar, etc.) et les bases de données contenant leur texte intégral (sources externes ou interfaces natives des fournisseurs par exemple). Sur swisscovery, lorsqu'un article est disponible en ligne, ce qui signifie qu'il fait partie des ressources souscrites ou qu'il est en

Open Access, un ou plusieurs de ces liens (*Services*) s'affichent dans la partie « Consulter en ligne » (*View It*) de la page de visualisation des notices (*Full Display*). Le résolveur de liens est donc central dans la compréhension de l'utilisation des ressources électroniques par les utilisateurs-trices de swisscovery.

Le rapport *Ressource Sharing Borrowing Requests_Per Partner*, qui comptabilise le nombre de documents demandés par la communauté HES-SO auprès d'autres bibliothèques, n'est disponible que depuis novembre 2022 et n'a donc pas pu être exploité.

3.3.3. Données COUNTER

Enfin le dernier corpus de données concerne les statistiques d'utilisation des ressources électroniques, issues des fournisseurs. Alma Analytics offre la possibilité d'éditer directement les rapports COUNTER depuis son interface, à la condition que les fournisseurs y soient paramétrés, ce qui n'est pas le cas de la HES-SO. Nous avons donc utilisé les rapports COUNTER directement fournis par les éditeurs.

Le code de pratique COUNTER standardise la production de statistiques d'usage liées à la consultation des ressources électroniques sur les plateformes fournissant des contenus (Mellins-Cohen, 2021, p. 3). Cependant, comme tous les fournisseurs ne mettent pas à disposition leurs données selon ce standard, il n'est pas possible de consolider l'ensemble des accès aux ressources électroniques par la communauté de la HES-SO.

Dans le cadre de notre recherche, nous avons exploité les rapports d'usage des journaux de deux des plus importants fournisseurs de contenus électroniques à la HES-SO : Elsevier et Springer. Ces rapports sont standardisés sous les noms TR-J1 et TR-J3 (Mellins-Cohen, 2021, p. 14) [Tableau 3].

Tableau 3 : Rapports COUNTER de trois fournisseurs.

Rapports COUNTER
1. Rapport TR-J1 d'Elsevier
2. Rapport TR-J3 * d'Elsevier
3. Rapport TR-J1 de Springer
4. Rapport TR-J3 * de Springer

Nous renvoyons à notre mémoire de recherche pour une présentation détaillée de l'ensemble de ces rapports (Kerboul & Leboulleux, 2024a, p. 23-50), en complément de la documentation d'Ex Libris disponible en ligne sur Primo VE Analytics (Ex Libris, 2022) et Alma Analytics (Ex Libris, 2023c), ainsi que du manuel de bonnes pratiques COUNTER (2023).

4. Résultats

4.1. Production des indicateurs

La structure de notre diagramme de Sankey est inspirée des tâches utilisateur génériques définies dans le modèle IFLA LRM, segmentant le processus de recherche documentaire en cinq grandes étapes (Riva et al., 2021, p. 15-17) :

1. « Trouver » des ressources répondant à certains critères de recherche
2. « Identifier » les ressources en les différenciant entre elles

3. « Sélectionner » en ne conservant que les ressources pertinentes
4. « Obtenir » la ressource en accédant à son contenu
5. « Explorer » pour découvrir d'autres ressources liées.

Ces tâches utilisateur sont caractéristiques des fonctionnalités attendues par les utilisatrices de toute plateforme de recherche documentaire, dont les outils de découverte, et permettent d'en cerner l'utilisation et la performance.

Notre parcours utilisateur reprend donc ce séquençage chronologique et en quantifie les différentes étapes à l'aide des données relatives aux domaines suivants :

1. Sessions
2. Recherches (« Trouver »)
3. Résultats (« Identifier »)
4. Visualisations (« Sélectionner »)
5. Accès (« Obtenir »)

Bien que le document de l'IFLA relève le caractère itératif du processus de recherche d'informations (Riva et al., 2021, p. 15), nous avons présenté ce parcours sous forme linéaire. La tâche « Explorer », pouvant intervenir à n'importe quel moment de la recherche documentaire et constituant de ce fait un cas particulier (Riva et al., 2021, p. 15), a fait l'objet d'analyses complémentaires. Nous renvoyons à notre mémoire de recherche sur ce point (Kerboul & Leboulleux, 2024a, p. 82-94).

4.1.1. Sessions

Sur Analytics, nous disposons de trois rapports dénombrant un nombre total de sessions sur Primo VE : *Actions per session*, *Device Usage* et *Session*. Or ces rapports font état de trois valeurs différentes pour 2022. En effet, les deux premiers rapports sont construits de telle façon qu'une même session peut être comptabilisée plusieurs fois.

C'est pourquoi nous avons utilisé la mesure *Sessions*, la plus faible des trois, issue du rapport éponyme *Session*, et qui, d'après la documentation d'Ex Libris, répond à la question « Combien de sessions ont été ouvertes au cours d'une période donnée ? » (Ex Libris, 2022, notre traduction).

Le nombre total de sessions sur swisscovery à la HES-SO en 2022 retenu, est donc de 388'401 sessions (mesure *Sessions* du rapport *Session*).

Rapporté au mois et à la population cible [Chapitre 3.2], nous pouvons estimer qu'en moyenne sur l'année et indépendamment de toute saisonnalité, chaque utilisateur-trice potentiel-le de swisscovery à la HES-SO a ouvert 1,5 sessions par mois. Cet indicateur présente néanmoins une forte variabilité mensuelle (écart-type de 0,50), étant donné le rythme académique bi-semestriel des études universitaires menées par une grande partie de la population cible (pics d'activité en mars, octobre et novembre et périodes creuses en juillet et août sur 2022).

De la même façon, et afin de rester cohérents, nous nous sommes référés au nombre d'actions issu du rapport *Session* dans le cadre de notre travail. Le ratio *Actions / Sessions* est ainsi de 3,4 sur 2022.

Enfin, la moyenne des durées de session, pondérée par le nombre de sessions de chaque pays, est de 2,9 minutes en 2022. Elle correspond au temps écoulé entre la première et la dernière actions enregistrées pour chaque session.

4.1.2. Recherches

Les données d'usage sur le nombre de recherches effectuées par les utilisateurs-trices de swisscovery à la HES-SO sont simples à compiler car elles sont toutes issues du rapport standard *Action Usage Searches*.

Nous conservons les différents types de recherches, afin de relater au mieux les pratiques informationnelles de la population, soit les recherches basiques (*Basic search*), les recherches avancées (*Advanced search*), les recherches vocales (*Voice Search*), les recherches depuis l'historique (*Session query*) et les recherches sauvegardées (*Saved query*), pour un total de 384'922 recherches, tous types confondus.

À ce stade, nous constatons que le nombre moyen de recherches par session en 2022 est de 1,0 (Recherches totales / Sessions).

4.1.3. Résultats

Étape pourtant clé de la recherche documentaire, nous n'avons que peu d'informations quantitatives sur les résultats issus des requêtes effectuées. De plus, le nombre de recherches sans résultat ne peut pas être obtenu avec le rapport *Zero Result Searches* pour deux raisons :

- Le nombre de recherches distinctes rapporté est plafonné à 100 par jour. Avec cinq occurrences en 2022, cette limitation n'a cependant pas beaucoup d'impact.
- Le rapport ne fournit que le nombre de recherches sans résultat distinctes sur une même journée et non le nombre total de recherches sans résultat.

Nous pouvons donc uniquement conclure que le nombre total de recherches sans résultat est au minimum de 19'328, correspondant au nombre de lignes du rapport *Zero Result Searches*. Ceci signifie qu'au minimum 5% des recherches totales effectuées en 2022 n'ont pas obtenu de résultat.

4.1.4. Visualisations

Nous entendons par « visualisation » l'affichage de la notice d'un document donné (*Full Display*). Cette visualisation est en grande majorité la conséquence d'un clic sur un des documents de la liste des résultats (*Brief Results*). Cette action est comptabilisée dans la mesure *Display full record* du rapport *Action Usage Document*. Le clic peut être réalisé n'importe où sur le résultat, à l'exception de la mention « Disponible en ligne » qui incrémente une mesure spécifique (*Click on availability statement*), sans impacter la valeur de *Display full record*. Nous devrions donc très logiquement ajouter cette mesure à nos calculs de visualisations. Or, il est également possible de cliquer sur « Disponible en ligne » depuis la notice elle-même (*Full Display*), ce qui ne correspond pas dans ce cas à une nouvelle visualisation. Face à l'impossibilité de dissocier ces deux actions (Ex Libris, 2023a), nous avons décidé de ne pas intégrer cette mesure.

Deux autres mesures nous ont interrogés (*Click on title* et *Click on icon*). La documentation ne précisant rien les concernant, nous sommes partis de l'hypothèse qu'ils étaient déjà pris en compte dans la mesure *Display full record*.

Enfin, nous avons identifié quatre autres actions permettant de visualiser un document en *Full Display* et que nous avons prises en compte :

- *Next Result* et *Previous Result*, du rapport *Action Usage Document*.
- *Click on Record from Virtual Browse* et *Click on record from More From The Same*, du rapport *Action Usage Exploration*.

Une dernière inconnue réside dans les potentielles autres façons d'accéder à l'affichage d'un document, depuis *Mes favoris* ou via un permalien par exemple, mais nous n'avons identifié aucune mesure semblant s'y rapporter.

En résumé, nous avons décidé de consolider ce que nous appelons « nombre de visualisations » en sommant les champs suivants : *Display full record*, *Next Result*, *Previous Result*, *Click on Record from Virtual Browse*, *Click on record from More From The Same*. Nous avons exclu les *Click on title* et les *Click on icon*, a priori à juste titre. Enfin, nous avons exclu les *Click on availability statement*, ce qui n'est que partiellement juste.

En l'état actuel de nos connaissances, nous retiendrons donc le nombre de 389'787 visualisations en 2022, soit 1,0 visualisation en moyenne par session.

Bien que peu probable, nous relevons néanmoins que cette valeur pourrait atteindre un maximum de 447'620 visualisations en 2022, en intégrant la totalité des *Click on availability statement*, soit 1,15 visualisations en moyenne par session.

4.1.5. Accès

La dernière étape du parcours utilisateur est l'accès aux ressources, tant physiques qu'électroniques (tâche utilisateur « Obtenir » de l'IFLA). Nous disposons tout d'abord, pour cette partie, de deux rapports issus d'Alma Analytics.

Prêts physiques

Les données nécessaires sont extraites du rapport *Count of Items Loaned by Patron Group*. Ce rapport comptabilise les prêts effectués par les usagers-ères des bibliothèques (*Not In House*) ainsi que les ouvrages utilisés dans l'enceinte de la bibliothèque, sans avoir été empruntés (*In House*).

Le nombre de prêts total, enregistrés sur Alma en 2022 dans les bibliothèques HES-SO membres du réseau swisscovery, est de 50'318 prêts (*Grand Total* du rapport *Count of Items Loaned by Patron Group*)

Accès aux ressources électroniques via swisscovery

Nous avons ensuite exploité le rapport *7a.Resolver Statistics by Source*, sur le résolveur de liens Alma, pour quantifier les accès au texte intégral des ressources électroniques via swisscovery.

Nous avons retenu les deux champs suivants pour la construction de notre parcours utilisateur :

- Le nombre total de fois où l'utilisateur-trice a effectivement tenté d'accéder au texte intégral de la ressource électronique visualisée, en cliquant sur au moins un des liens proposés sur la notice (*Number of Clicked Requests*)

- Le nombre total de fois où la requête OpenURL n'a abouti à aucun service et où l'utilisateur-trice, indépendamment de sa volonté supposée de cliquer sur l'accès au texte intégral de la ressource, n'en a de toute façon pas eu la possibilité (*Number of Requests Without Services*).

Ainsi, le nombre total de ressources électroniques visualisées, pour lesquelles au moins un clic pour accéder au texte intégral a été enregistré, s'élève à 65'311 en 2022 (*Number of Clicked Requests*).

En parallèle, le nombre total de ressources électroniques pour lesquelles l'accès était impossible lors de leur visualisation est de 66'837 en 2022 (*Number of Requests Without Services*).

La difficulté à laquelle nous nous sommes heurtés par la suite a résidé dans l'analyse plus fine de ces données, au niveau des fournisseurs.

Nous avons en effet besoin d'extraire de ce rapport les données relatives à l'accès aux journaux des deux fournisseurs retenus pour notre étude, à savoir Elsevier et Springer, pour être en mesure de les comparer à leurs statistiques COUNTER.

Pour atteindre cet objectif, nous avons été contraints à certaines approximations qui ont plutôt maximisé le nombre d'accès aux ressources électroniques de ces deux fournisseurs (Kerboul & Leboulleux, 2024a, p. 79-82).

Ainsi, parmi les 65'311 accès aux ressources électroniques depuis swisscovery HES-SO en 2022, nous estimons qu'au maximum 10'202 sont issues des collections de journaux d'Elsevier et 2'492 des collections de journaux de Springer.

Toutes les valeurs nécessaires à la construction du parcours utilisateur dans swisscovery sont désormais calculées.

Accès global aux ressources électroniques de Elsevier et Springer

La dernière étape de notre analyse de données consiste à estimer le volume total de ressources électroniques consultées pour Elsevier et Springer par la communauté de la HES-SO en 2022, à l'aide des données COUNTER de ces deux fournisseurs.

Dans ce but :

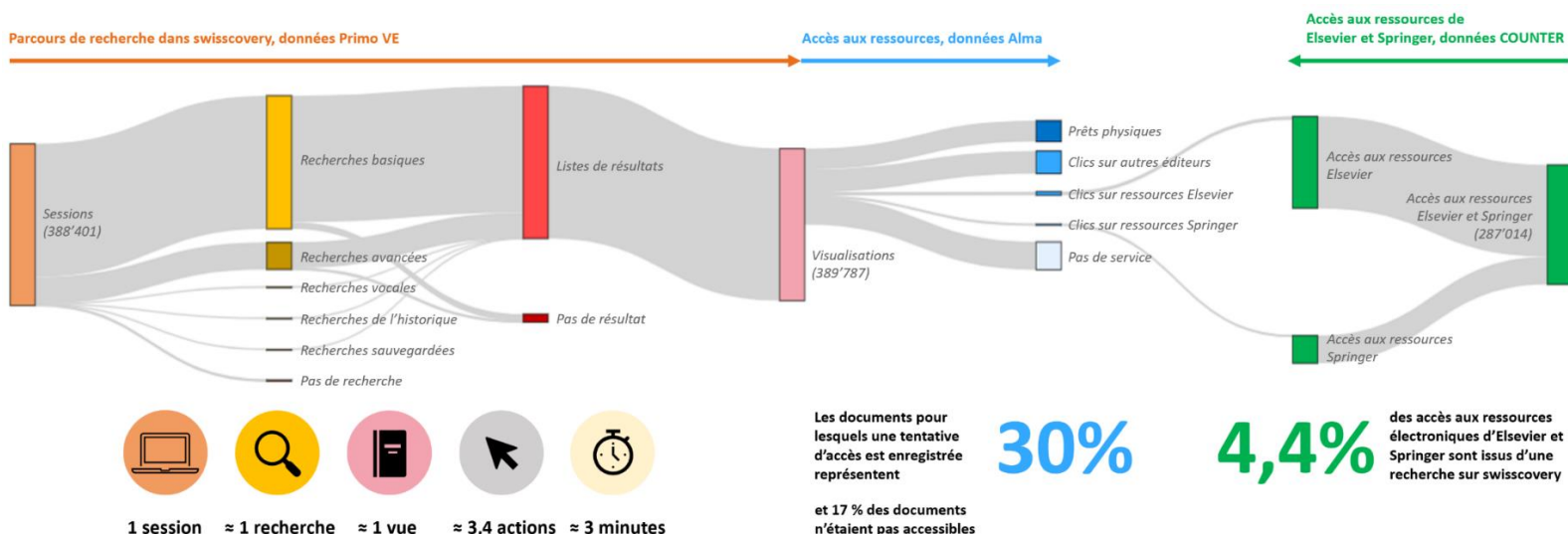
- Nous avons privilégié le rapport TR-J3 à TR-J1, car il intègre les ressources en Open Access (Mellins-Cohen, 2021, p. 14), ce qui correspond au même périmètre que le rapport *7a.Resolver Statistics by Source* de Alma Analytics
- Nous avons choisi d'utiliser la mesure *Unique_Item_Requests*, et non pas *Total_Item_Requests*, car elle ne comptabilise que le nombre de documents uniques dont un utilisateur-trice a demandé le texte intégral au cours d'une même session (COUNTER, 2023, p. 29).

En conclusion, en effectuant le total des *Unique_Item_Requests* des rapports TR-J3 d'Elsevier et de Springer, nous obtenons un total de 287'014 accès au texte intégral, dont 220'634 pour Elsevier et 66'380 pour Springer.

4.2. Diagramme de Sankey du parcours utilisateur

L'ensemble des indicateurs identifiés, définis et calculés précédemment est intégré à la visualisation d'un parcours utilisateur type dans swisscovery HES-SO, sous la forme d'un diagramme de Sankey, généré dans R [Figure 1].

Figure 1 : Diagramme de Sankey du parcours de recherche documentaire dans swisscovery HES-SO en 2022.



Notre diagramme adopte une perspective similaire à l'étude de Jiang et al. (2017, p. 220), mais inclut des données de sources hétérogènes, issues de Primo VE (en orange), d'Alma (en bleu) et de COUNTER (en vert), afin de pouvoir reconstituer un parcours complet sur la base des données disponibles. Cette variété nous a contraints à certaines approximations, afin d'être en mesure de lier les différents jeux de données entre eux. Ainsi les prêts physiques représentent des maxima, car dans les faits tous les prêts physiques ne résultent pas d'une recherche sur swisscovery. Ce diagramme fait néanmoins la synthèse de nombreux indicateurs et présente l'avantage d'être synthétique, visuel et facilement compréhensible. En ce sens, il se rapproche d'un tableau de bord.

Il met en évidence l'utilisation standard qui est faite de swisscovery, et peut se lire aussi bien de gauche à droite que de droite à gauche. La lecture de gauche à droite suit le processus de recherche documentaire. Elle permet de visualiser la répartition par type de recherche et le taux de recherche sans résultat. L'homogénéité de la largeur du ruban sur l'ensemble du processus met en évidence la linéarité des sessions, correspondant généralement à une recherche et une visualisation. Ce diagramme interroge cependant sur l'après visualisation, où nous n'avons pu identifier que 47% du chemin subséquent.

La lecture de droite à gauche permet quant à elle de replacer l'usage de swisscovery dans son écosystème. En effet le diagramme met en évidence que le nombre total d'accès aux ressources électroniques (ici pour Elsevier et Springer) est bien supérieur au trafic sortant de swisscovery pour atteindre ces mêmes ressources électroniques. Il interroge donc sur la place

de swisscovery dans les pratiques informationnelles de la population cible de la HES-SO et sur les parcours alternatifs utilisés.

Enfin, nous avons décidé de mettre en avant les indicateurs généraux permettant de qualifier une session type, à savoir qu'une session en 2022 correspond en moyenne à une recherche et une visualisation de document, engendre 3,4 actions et dure environ 3 minutes.

Ces indicateurs permettent d'appréhender l'usage type moyen qui est fait de la plateforme swisscovery : un outil de recherche documentaire pour accéder à une ressource précise, rapidement et en peu d'actions.

5. Discussion

5.1. Interprétation des résultats

Les résultats et les conclusions suivantes peuvent être avancées sur la base de notre étude, concernant l'usage de swisscovery HES-SO pour l'année 2022 :

- **Le nombre de sessions, rapporté à sa population cible et par mois, est de 1,5 sessions par personne et par mois.**

Ce ratio pourra être comparé avec les données des prochaines années, dans le but d'en évaluer l'évolution. Néanmoins, nous savons par l'analyse des prêts que les utilisateurs-trices de swisscovery HES-SO ne se limitent de loin pas à la seule communauté académique de la HES-SO (Kerboul & Leboulleux, 2024a, p. 67).

- **En moyenne, une session correspond à une recherche et une visualisation.**

Le fait que des sessions ne soient pas ouvertes à des fins de recherche documentaire, mais pour gérer uniquement son compte ou ses prêts, aurait pu influencer sur ces résultats. Mais les données d'usage relatives à ces actions demeurent suffisamment faibles (13'629 accès aux prêts, dont 6'313 prolongations, et 8'376 demandes en 2022 par exemple), pour que leur impact ne soit pas jugé significatif.

Parmi les résultats d'autres études, Jiang et al. (2017, p. 219) rapportent 1,42 recherches et 0,57 visualisations par session. De leur côté, 80% des visiteurs-euses de Gallica ne consultent qu'un seul document, tandis que 10 % en visualisent 5 (Brandl, 2017, p. 7).

Ces valeurs laissent à penser que les utilisateurs-trices lancent une recherche pour accéder à un document précis, ou alors se contentent d'un seul résultat.

- **En moyenne, une session engendre 3,4 actions.**

Les usagers-ères de swisscovery ont une utilisation très simple de la plateforme, qui se résume à quelques actions seulement, telles que consulter son compte et prolonger ses prêts, ou, en très grande majorité, faire une seule recherche pour n'en consulter que quelques résultats. Jiang et al. (2017, p. 219) rapportent des conclusions similaires.

- **En moyenne, une session dure 2,9 minutes.**

Jiang et al. (2017, p. 219) rapportent un temps moyen supérieur à 5,3 minutes par session mais arrivent à la même conclusion : « users wanted to find a couple of known resources rather than discovering possibly useful or interesting ones via immersive exploration ». Huttenlock et Malone (2013, p. 376) quant à eux ne calculent pas la durée moyenne mais rapportent que 30% des utilisateurs-trices restent moins d'une minute et 30 autres pourcents moins de cinq minutes. Concernant la bibliothèque numérique Gallica, Brandl (2017, p. 6) relève des durées de session majoritairement inférieures à 2 minutes, qualifiées d'« assez courtes ».

Cette durée moyenne semble effectivement trop courte pour initier une démarche exploratoire sur un outil de découverte, supposé encourager la navigation parmi les ressources.

- **Les utilisateurs-trices privilégient très majoritairement la recherche basique (83%) à la recherche avancée (17%).**

Cohen et Thorpe (2015, p. 39) quant à elles rapportent qu'entre 11% et 27% des recherches étaient faites en mode avancé.

Les utilisateurs-trices ont généralement un usage basique de la plateforme, et interrogent majoritairement swisscovery de la même façon qu'un moteur de recherche de type Google, en saisissant leurs mots clés dans la barre de recherche proposée, tous champs confondus.

- **Le taux de recherche sans résultat est de 5% au minimum.**

En comparaison, Huttenlock et Malone (2013, p. 376) rapportent que 4% des recherches n'ont pas abouti.

Une analyse qualitative des principales causes d'échec et des réactions des utilisateurs-trices est disponible dans notre mémoire. Nous avons en outre identifié deux recherches enregistrées comme sans résultat alors que, selon nous, elles auraient dû en fournir (Kerboul & Leboulleux, 2024a, p. 87-89).

- **30% maximum des documents visualisés font l'objet d'une tentative d'accès (prêt physique ou accès au texte intégral).**

En pratique, les visualisations sans accès peuvent avoir de nombreuses causes : la ressource consultée s'avère non pertinente, la consultation avait pour but unique de récupérer une référence bibliographique, de localiser le document physique en rayon à partir de sa cote (sans qu'il soit comptabilisé en prêt *In House* par la suite) ou d'en faire une demande de numérisation, le document physique n'est pas disponible ou l'utilisateur-trice renonce à le demander s'il n'est pas en libre accès, etc. Il peut également s'agir de consultations de notice par le personnel bibliothécaire, pour leurs besoins propres ou pour renseigner des usagers-ères.

- **17% des documents visualisés sont des ressources électroniques dont l'accès était impossible (pas de lien proposé sur la notice).**

Si l'on ne considère que les ressources électroniques, le taux de requêtes OpenURL sans service atteint 21%.

Cette valeur représente un dysfonctionnement dans l'accessibilité des ressources. En l'état, le périmètre de notre revue de littérature ne nous a pas permis de comparer ce résultat avec d'autres études.

- **Seuls 4,4% des accès aux journaux électroniques de Elsevier et Springer sont issus d'une recherche sur swisscovery.**

Ce résultat met en évidence l'existence d'autres parcours informationnels au sein de la communauté de la HES-SO et démontre que les journaux de ces deux fournisseurs sont en très grande majorité accédés par d'autres moyens que par une recherche documentaire dans swisscovery.

Parmi les autres parcours possibles, nous avons connaissance de :

- Google Scholar, moteur de recherche d'une très grande facilité d'utilisation et d'un temps de réponse très court.
- Les plateformes des fournisseurs de contenus (ProQuest, Web of Science, DOAJ, etc.), les archives de publications scientifiques et dépôts institutionnels (JSTOR, HAL, etc.), et les bases de données spécialisées (MEDLINE via PubMed, LISA,

etc.). L'utilisateur-trice se retrouve très souvent sur l'une de ces interfaces à l'issue d'une première recherche sur swisscovery et peut tout à fait y poursuivre ses investigations. L'ergonomie et la performance de swisscovery sont donc en concurrence directe et permanente avec celles de ces nombreuses plateformes.

- Les extensions de navigateur. Par exemple « Click & Read » de l'Inist CNRS signale les ressources électroniques disponibles, en Open Access ou souscrites pour les organismes partenaires, en affichant un logo sur les pages web ou documents html consultés, et donne immédiatement accès au texte intégral, sans aucune recherche préalable.
- Dans le cas de la HES-SO, les utilisateurs-trices du réseau Renouvaud et des catalogues indépendants des bibliothèques non-membres de swisscovery.

Sur ce point, la littérature scientifique fait allusion de façon récurrente à la place très minoritaire des outils de découverte dans l'accès aux ressources électroniques (Raieli, 2022, p. 38; Robins, 2022), mais souvent sans avancer de chiffres précis. Nous pouvons néanmoins évoquer Cummings (2021, p. 69), qui avance la valeur moyenne de 10% des HTTP Referers en provenance de l'outil de découverte, variant de 3,2% à 17,8% selon les éditeurs.

5.2. Limitations sur les données

Notre étude a veillé à s'assurer que les indicateurs produits correspondent à la réalité des phénomènes dont nous souhaitons rendre compte et qu'ils soient suffisamment documentés pour être interprétés correctement par la suite, en tenant compte de leur périmètre et de leurs limites.

La validité des résultats obtenus est néanmoins fortement conditionnée par la qualité des données dont ils sont issus, tant sur leur accessibilité, exploitabilité ou reproductibilité, ainsi que par le niveau de compréhension des gestionnaires de ces données.

Forts de ce constat, nous relevons un certain nombre de limites à notre recherche, sur la base desquelles nous pouvons émettre plusieurs recommandations.

• Un accès restreint aux données

Nous avons tout d'abord été confrontés à une difficulté d'accès aux données. Ne disposant pas d'un compte sur Primo VE et Alma Analytics, nous avons dû solliciter le CISO pour réaliser les extractions nécessaires à notre recherche. À ce titre, nous remercions les collaboratrices du CISO pour le temps accordé.

Toutefois, nous pensons que cette absence d'accès direct a eu des impacts négatifs, en limitant nos possibilités d'exploration, à la fois des données et des fonctionnalités d'Analytics. Dans l'hypothèse d'autres recherches de ce type, menées par des étudiant-e-s, nous recommandons vivement d'investiguer plus en avant la possibilité de leur ouvrir un accès direct à Analytics, encadré par des clauses de confidentialité des données si nécessaire.

• Des données complexes à consolider

Nous avons sous-estimé la complexité engendrée par la coexistence des réseaux swisscovery et Renouvaud, ainsi que de bibliothèques disposant d'un catalogue indépendant au sein de la HES-SO. Indépendamment de toute autre considération, et perçu uniquement du point de vue de la gestion des données, il nous semblerait tout à fait souhaitable de pouvoir consolider les données de l'ensemble de la HES-SO via un seul et même outil de gestion.

- **Des données partielles**

Certaines données se sont avérées partielles. Ainsi le rapport *Primo VE Zero Result Searches* ne fournit pas le nombre total de recherches sans résultat, et nous a contraint à en estimer un nombre minimal sur une période donnée. Pour ce rapport en particulier, nous recommandons de tester la possibilité d'ajouter le nombre total de recherches effectuées, telle que la mesure *Searches* présente dans le rapport *Popular Searches*.

- **Des données trop agrégées**

Par leur construction, les rapports standards fournissent souvent des données très agrégées, restreignant ainsi les possibilités d'analyse. Les données issues de la table *Action Usage* de Primo VE, par exemple, ne sont extraites qu'avec un détail mensuel. De façon générale, il conviendrait d'explorer les possibilités de personnalisation des rapports standards pour en augmenter les axes d'analyse potentiels et pouvoir opérer, par exemple, la distinction entre les différentes catégories d'utilisateurs-trices (étudiant-e-s ou professeur-e-s, par filière, par école, etc.).

- **Des données mal documentées**

Nous nous sommes également confrontés à la documentation des modules Analytics de Primo VE et d'Alma, mais aussi des logiciels eux-mêmes, disponibles sur le site *Ex Libris Knowledge Center* (Ex Libris, 2023b). Le fait que la documentation de Primo VE ne soit pas disponible en français nous a contraints à de nombreux aller-retours entre les interfaces en français et en anglais de swisscovery, pour identifier ce à quoi pouvaient correspondre les différentes actions possibles. Nous nous sommes également parfois retrouvés assez démunis face à des définitions de champ très succinctes et peu explicites, ou au contraire à des notes dont nous ne saisissons pas toute la portée. Nous avons partagé le même constat lors d'échanges informels avec des professionnels : la documentation actuelle ne constitue pas la plus simple des entrées en matière pour s'approprier les modules Analytics de Primo VE et Alma et ce d'autant plus que nous ne pouvions pas faire de tests sur Analytics pour valider nos hypothèses. Le CISO a par ailleurs été amené à interroger la communauté Ex Libris pour notre recherche, mais la réponse s'est avérée trop technique pour répondre à notre besoin.

- Dans ce contexte, nous alertons sur le fait que notre niveau de compréhension s'est limité à ce que nous avons pu assimiler dans le temps imparti et aux moyens mis à notre disposition.

5.3. Limitations méthodologiques

Nous apportons également un regard critique quant à notre méthodologie.

Nous n'avons pas été en mesure d'identifier de données relatives à l'accès aux documents, que ce soit pour le prêt ou vers le texte intégral des ressources électroniques, directement de Primo VE Analytics. Cela nous a contraints à exploiter les statistiques de prêts et le rapport sur le résolveur de liens issus d'Alma. D'une part les prêts totaux ne sont pas directement liés aux recherches effectuées dans swisscovery et d'autre part nous avons manqué de documentation pour le traitement des données du résolveur de liens. Enfin, notre analyse des données COUNTER s'est limitée à deux fournisseurs. Il conviendrait d'en explorer d'autres pour vérifier si les ratios obtenus sont du même ordre de grandeur ou si l'accès à des fournisseurs plus spécialisés est plus élevé via swisscovery, leurs collections étant rendues plus visibles par exemple.

De façon générale, nous ne comprenons pas pourquoi Primo VE Analytics offre autant de données sur des actions de moindre intérêt, telles que la gestion de ses favoris ou les types

de citations demandés, mais soit si obscur sur les données essentielles que sont les accès effectifs aux documents. Nous ne pouvons que souhaiter une meilleure clarté sur ce point.

Ces trois sources de données distinctes ont néanmoins permis de recomposer un parcours type dans swisscovery, qui a l'intérêt d'offrir une compréhension aisée. Le diagramme de Sankey présente toutefois lui aussi des approximations, car nous ne disposons pas de toutes les données requises pour sa construction. Ainsi, les sessions sont réparties sur l'ensemble des recherches, mais cela ne signifie pas que toutes ces sessions donnent lieu à des recherches. En pratique, une même session peut en effet donner lieu à plusieurs recherches et le nombre réel de sessions sans recherche est donc probablement plus élevé que visualisé. De même, nous n'avons pas l'information sur le taux de recherches sans résultat à l'issue de chaque type de recherche et avons donc dû répartir ces recherches sans résultat de façon uniforme sur les recherches basiques et avancées.

5.4. Recommandations

5.4.1. Des modules Analytics à investir

Il ressort déjà des éléments de discussion avancés que l'appropriation complète des modules Analytics de Primo VE et Alma constitue un enjeu central pour garantir la compréhension et assurer le suivi de l'usage de swisscovery.

Analytics permet par exemple de personnaliser les rapports standards, en augmentant la granularité des données, temporelle par exemple, ou en ajoutant des axes d'analyse tels que *Institution*, *User Group* ou *Primo View*. Ce sont autant de nouvelles possibilités d'analyse qu'il ne nous a pas été possible d'explorer, sans compter que des rapports « sur mesure » pourraient également être créés ex nihilo, pour alimenter des tableaux de bord.

Toutefois, à notre connaissance, peu de personnes semblent s'être appropriées toutes les fonctionnalités d'Analytics et avoir les compétences pour en exploiter tout le potentiel. L'expertise technique appartient à SLSP, auprès de qui il est possible de faire des demandes via une ouverture de ticket, mais aucune formation spécifique à Analytics ne semble exister. Il ne semble pas non plus exister de procédures visant à uniformiser la production des données et les rapports semblent difficiles à retrouver sur l'interface.

Moyennant un investissement non négligeable en temps, la connaissance approfondie de Analytics et des données d'usage collectées constituerait un réel avantage pour la HES-SO.

5.4.2. Des axes d'étude complémentaires à investiguer

Notre recherche a exclu deux axes d'étude qui ouvrent la porte à de nouveaux travaux.

Le premier consiste à recenser les pratiques de suivi des données d'usage établies dans d'autres institutions académiques, telles que l'UNIGE, l'EPFL, l'ETH, ou même la BCUL sur Renouvaud, afin d'en extraire de bonnes pratiques généralisables.

Le second vise l'exploitation des fichiers journaux (*logfiles*), si disponibles, dans le but d'affiner le degré d'analyse et de compréhension du parcours utilisateur.

6. Conclusion

Notre objectif était de comprendre l'usage de la plateforme swisscovery à la HES-SO, en se basant sur des données quantitatives. À l'issue de notre travail, nous sommes en mesure de conclure que celle-ci interroge sur plusieurs points.

Tout d'abord, les données issues de Primo VE Analytics nous ont appris qu'une session moyenne se résumait à 3-4 clics sur l'interface, effectués en 3 minutes, correspondant à une recherche, généralement basique, suivie d'une seule visualisation de document. Swisscovery répond donc majoritairement à un besoin informationnel précis, visant une ressource en particulier. Ce constat paraît éloigné des buts d'un d'outil de découverte, favorisant l'exploration des collections par navigation. Nous rejoignons ainsi les conclusions de Jiang et al. (2017, p. 219).

Les données extraites d'Alma Analytics nous ont ensuite renseignés sur le taux d'accès effectif aux ressources visualisées sur swisscovery, c'est-à-dire un prêt physique ou une consultation du texte intégral numérique. Ce taux était de maximum 30% en 2022 (en comptabilisant l'ensemble des prêts physiques), auxquels s'ajoutent 17% de ressources électroniques ne proposant pas de service d'accès. À l'issue de notre projet de recherche, ce taux d'accès aux documents, que l'on peut considérer comme un taux d'engagement des utilisateurs-trices, nous avait semblé faible. Désormais, il nous semble plutôt caractéristique de recherches ciblées, ce qui est cohérent avec des sessions de courte durée.

Enfin, la mise en perspective des accès aux ressources électroniques comptabilisés par le résolveur de liens d'Alma, avec les données COUNTER fournies par deux éditeurs, permet d'avancer que seuls 4,4% des accès totaux aux ressources de ces fournisseurs étaient issus d'une recherche sur swisscovery en 2022. Ce faible taux témoigne d'autres parcours informationnels pratiqués par la population cible de swisscovery à la HES-SO, ainsi que d'une forte concurrence d'autres canaux d'accès aux ressources électroniques.

À court terme, ce constat, qui relativise l'importance de swisscovery dans le processus de recherche documentaire de la communauté académique de la HES-SO, incite à prioriser l'accessibilité des ressources électroniques, par leur mise à disposition correcte et sans défaut de service, plutôt que leur visibilité dans swisscovery.

À moyen terme, nous encourageons les institutions du réseau swisscovery à échanger sur ces indicateurs, issus de leurs données d'usage, dans le but de consolider les résultats obtenus voire d'en étendre la portée.

À plus long terme, et sur le périmètre des ressources électroniques uniquement, il conviendrait de se questionner sur l'avenir des outils de découverte face à la montée en puissance de l'Open Access, et à la concurrence accrue des moteurs de recherche et des bases de données spécialisées.

À ce sujet, S. Kouassi Kouakou (2021) nous invite à un changement paradigmatique, afin de « passer de la logique de stockage de documents à celle de milieu de production du savoir », s'appliquant également aux catalogues :

« augmentés avec des notices intelligentes qui agrègent un ensemble d'information descriptive, analytique et cartographique sur les documents, [...] »

pour assurer une médiation renouvelée fondée sur une relation plus directe et plus vivante avec les collections. » (Kouakou, 2021, p. 87).

Bibliographie

- Alotaibi, F. A. A., Johnson, F., & Rowley, J. (2022). Google Scholar or University Digital Libraries : A comparison of student perceptions and intention to use. *Journal of Librarianship and Information Science*, 55(4), 906-920. <https://doi.org/10.1177/09610006221111197>
- Brandl, E. (2017). Enseignements et perspectives à partir de Gallica. Journée d'étude du 3 mai 2017. *Les billets d'EnssibLab*. <https://www.enssib.fr/bibliotheque-numerique/documents/68776-diversite-des-usages-des-bibliotheques-numeriques-methodes-d-etude-a-partir-de-l-exemple-de-gallica.pdf>
- Breeding, M. (2009). Next-Generation Flavor in Integrated Online Catalogs. *Library Technology Reports*, 43(4), Article 4.
- Bullock, C. (2021). Finding Open Content in the Library Is Surprisingly Hard. *Serials Review*, 47(2), 68-70. <https://doi.org/10.1080/00987913.2021.1936416>
- Cohen, R. A., & Thorpe, A. (2015). Discovering User Behavior : Applying Usage Statistics to Shape Frontline Services. *The Serials Librarian*, 69(1), 29-46. <https://doi.org/10.1080/0361526X.2015.1040194>
- COUNTER. (2023). *COUNTER Code of Practice Release 5.0.3*. https://cop5.projectcounter.org/_downloads/en/5.0.3/pdf/
- Cummings, J. (2021). Online Navigation to Journal Articles : How Are Journal Articles Retrieved by Researchers and Students at an Academic Institution; a Quantitative Examination of HTTP Referer [sic] Data. *Journal of Electronic Resources Librarianship*, 33(2), 63-74. <https://doi.org/10.1080/1941126X.2021.1911446>
- Durante, K., & Wang, Z. (2012). Creating an Actionable Assessment Framework for Discovery Services in Academic Libraries. *College & Undergraduate Libraries*, 19(2-4), 215-228. <https://doi.org/10.1080/10691316.2012.693358>
- Evans, G., & Schonfeld, R. C. (2020). *It's Not What Libraries Hold; It's Who Libraries Serve. Seeking a User-Centered Future for Academic Libraries*. Ithaca S+R. <https://doi.org/10.18665/sr.312608>
- Ex Libris. (2022, mars 24). *Primo Analytics Subject Areas*. Ex Libris Knowledge Center. https://knowledge.exlibrisgroup.com/Primo/Product_Documentation/Primo/Analytics/040Primo_Analytics_Subject_Areas
- Ex Libris. (2023a, juillet 20). *Analytics Subject Areas for Primo VE*. Ex Libris Knowledge Center. [https://knowledge.exlibrisgroup.com/Primo/Product_Documentation/020Primo_VE/Primo_VE_\(English\)/140Primo_VE_Analytics/040Analytics_Subject_Areas_for_Primo_VE](https://knowledge.exlibrisgroup.com/Primo/Product_Documentation/020Primo_VE/Primo_VE_(English)/140Primo_VE_Analytics/040Analytics_Subject_Areas_for_Primo_VE)

- Ex Libris. (2023b, août 15). *Knowledge Center*. Knowledge.Exlibrisgroup.Com. <https://knowledge.exlibrisgroup.com>
- Ex Libris. (2023c, novembre 19). *Alma Analytics Subject Areas*. Ex Libris Knowledge Center. [https://knowledge.exlibrisgroup.com/Alma/Product_Documentation/010Alma_Online_Help_\(English\)/080Analytics/Alma_Analytics_Subject_Areas](https://knowledge.exlibrisgroup.com/Alma/Product_Documentation/010Alma_Online_Help_(English)/080Analytics/Alma_Analytics_Subject_Areas)
- Gustafson-Sundell, N. (2023). *Journal Data : Usage*. Minnesota State University Mankato. <https://libguides.mnsu.edu/c.php?g=1287205&p=9451851>
- Heinemann, C., Unterholzner, M.-C., & Marty, T. (2023). How to Support Diverse Libraries in a National Network? *ABI Technik*, 43, 96-102. <https://doi.org/10.1515/abitech-2023-0017>
- HES-SO. (2021). *A propos du CISO*. HES-SO. <https://www.hes-so.ch/recherche-innovation/information-scientifique/a-propos-du-ciso>
- HES-SO. (2022). *Liste des bibliothèques HES-SO membres de Renouvaud*. <https://intra3.hes-so.ch/modules/InternetFiles/readfile.asp?SharepointDLCDocId=HESSO-895-10926>
- HES-SO. (2023). *Liste des bibliothèques HES-SO membres de swisscovery*. <https://intra3.hes-so.ch/modules/InternetFiles/readfile.asp?SharepointDLCDocId=HESSO-900-6127>
- Huttenlock, T., & Malone, D. B. (2013). Proving Value and Improving Practice : Using System Data to Analyze User Behaviors. *College & Undergraduate Libraries*, 20(3-4), 366-385. <https://doi.org/10.1080/10691316.2013.829381>
- ISO. (2014). *Information et documentation—Méthodes et procédures pour évaluer l'impact des bibliothèques* (ISO 16439:2014; Version 1). <https://www.iso.org/fr/standard/56756.html>
- ISO. (2022). *Information et documentation—Statistiques internationales de bibliothèques* (ISO 2789:2022; Version 6). <https://www.iso.org/fr/standard/78525.html>
- ISO. (2023). *Information et documentation—Indicateurs de performance des bibliothèques* (ISO 11620:2023; Version 4). <https://www.iso.org/fr/standard/83126.html>
- Jiang, T., Chi, Y., & Gao, H. (2017). A clickstream data analysis of Chinese academic library OPAC users' information behavior. *Library & Information Science Research*, 39(3), 213-223. <https://doi.org/10.1016/j.lisr.2017.07.004>
- Jouneau, T., & Porquet, T. (2019). Évaluer l'usage des ressources numériques : Quels enjeux, quelles solutions ? In C. Touitou (Éd.), *Évaluer la bibliothèque par les mesures d'impacts* (p. 140-149). Presses de l'ENSSIB. <https://doi.org/10.4000/books.pressesensib.5758>
- Kerboul, T., & Leboulleux, D. (2024a). Comprendre l'usage de swisscovery à la HES-SO par les données [Mémoire de recherche]. Haute école de gestion de Genève HEG-GE. <https://zenodo.org/records/10817823>
- Kerboul, T., & Leboulleux, D. (2024b). *Swisscovery HES-SO_Dataset_Primo VE_2022* [Excel ou CSV]. Zenodo. <https://doi.org/10.5281/zenodo.13732434>
- Kouakou, S. K. (2021). Humanités numériques et bibliothèques : Des services innovants, un catalogue augmenté et une médiation renouvelée. *Revue maghrébine de documentation et d'information*, 30, 69-89. <https://hal.science/hal-03573812>

- Mellins-Cohen, T. (2021). *Release 5.0.2: The Friendly Guide for Librarians*. COUNTER. <https://medialibrary.projectcounter.org/file/The-Friendly-Guide-for-Librarians>
- Merminod, M. (2023). Usages et utilisabilité d'un outil de découverte en contexte académique : Évaluation de l'expérience utilisateur de swisscovery UNIGE. *Revue électronique suisse de science de l'information (RESSI)*, 23, Article 23. <https://doi.org/10.55790/journals/ressi.2023.e1106>
- Moore, M. (2016). But is My Resource Included? How to Manage, Develop, and Think about the Content in Your Discovery Tool. *The Serials Librarian*, 70(1-4), 149-157. <https://doi.org/10.1080/0361526X.2016.1147877>
- Namei, E., & Young, C. A. (2015). Measuring Our Relevancy : Comparing Results in a Web-Scale Discovery Tool, Google & Google Scholar. *Association of College & Research Libraries (ACRL)*, 522-535. <http://hdl.handle.net/11213/17887>
- Ngo, L., Hennesy, C., & Knabe, I. (2019). The Impact of Web-Scale Discovery on the Use of Electronic Resources. *Serials Review*, 45(4), 227-238. <https://doi.org/10.1080/00987913.2019.1695343>
- Nicholas, D., Boukacem-Zeghmouri, C., Rodríguez-Bravo, B., Xu, J., Watkinson, A., Abrizah, A., Herman, E., & Świgoń, M. (2017). Where and how early career researchers find scholarly information. *Learned Publishing*, 30(1), 19-29. <https://doi.org/10.1002/leap.1087>
- Office fédéral de la statistique (OFS). (2023a). *Données de référence et définitions des variables* (do-f-16.02.02-2023-2). <https://www.bfs.admin.ch/asset/fr/24247269>
- Raieli, R. (2022). *Web-Scale Discovery Services : Principles, Applications, Discovery Tools and Development Hypotheses*. Elsevier Science. <https://hesge.scholarvox.com/catalog/book/docid/88930393>
- Riva, P., Le Boeuf, P., Žumer, M., Le Moullec-Rieu, A., Leresche, F., Puyrenier, F., & Roche, M. (2021). *IFLA LRM : Un modèle conceptuel pour l'information bibliographique* (Committee Publications). International Federation of Library Associations and Institutions (IFLA). <https://repository.ifla.org/handle/123456789/1703>
- Robins, S. (2022). Full-text retrievals and EBSCO Discovery Service : Assessing usage of e-journals across multiple platforms. *Journal of Electronic Resources Librarianship*, 34(2), 89-107. <https://doi.org/10.1080/1941126X.2022.2064105>
- SLSP. (2023). *Swiss Library Service Platform*. SLSP Swiss Library Service Platform. <https://slsp.ch>
- Wang, Y. (2020). Web-scale discovery and Google Scholar : A study in use patterns. *Journal of Electronic Resources Librarianship*, 32(1), 1-10. <https://doi.org/10.1080/1941126X.2019.1709722>
- Wang, Y., & Howard, P. (2012). Google Scholar Usage : An Academic Library's Experience. *Journal of Web Librarianship*, 6(2), 94-108. <https://doi.org/10.1080/19322909.2012.672067>
- Werlen, R. (2017). *Stratégie nationale suisse sur l'Open Access—Plan d'action* (p. 46). Swissuniversities. https://www.swissuniversities.ch/fileadmin/swissuniversities/Dokumente/Hochschulpolitik/Open_Access/Plan_d_action-f.pdf?sword_list%5B0%5D=CAS

Wu, M., & Chen, S. (2014). Graduate students appreciate Google Scholar, but still find use for libraries. *The Electronic Library*, 32(3), 375-389. <https://doi.org/10.1108/EL-08-2012-0102>