

e-rara.ch : une bibliothèque numérique pour les livres anciens

Alexis Rivier

alexis.rivier@ville-ge.ch

<https://orcid.org/0009-0004-8657-0852>

Bibliothèque de Genève

La bibliothèque numérique e-rara.ch est née dans le cadre plus général du programme d'investissement de la Bibliothèque électronique suisse, e-lib.ch, qui a financé pour la période 2008-2011 un ensemble de projets d'infrastructures pour l'information scientifique et technique⁽¹⁾. Elle offre aujourd'hui plus de 26'000 titres à la consultation et reçoit environ 70'000 visites par mois. Son rayonnement dépasse largement les frontières et les utilisateurs du portail situés aux États-Unis et en Allemagne sont même plus nombreux que ceux localisés en Suisse.

Désormais pérennisé, e-rara.ch est solidement implanté au sein de l'offre numérique suisse et poursuit sa croissance. Sept ans après ses débuts, un examen critique de la genèse de cette réalisation permet de montrer comment les problèmes techniques et organisationnels ont pu être surmontés.



Illustration 1. www.e-rara.ch: page d'accueil.

1. Le point de départ

Avant que le projet ne prenne forme en 2007, la perception générale en Suisse est que la numérisation des collections de bibliothèques n'a pas encore reçu l'attention qu'elle mérite en dehors des instances professionnelles. Les initiatives publiques piétinent, alors qu'ailleurs les grandes bibliothèques nationales passent d'une numérisation de niche à une numérisation de "masse". L'annonce du projet de Google, en décembre 2004, de numériser des collections intégrales en partenariat avec de prestigieuses bibliothèques anglo-saxonnes a fait l'effet d'un électrochoc. Marie-Christine Doffey, directrice de la Bibliothèque nationale suisse, constate avec

regret dans un article que "la numérisation des bibliothèques publiques avance à pas d'escargot " et qu'"une prise de conscience et une volonté politique font encore défaut en Suisse"⁽²⁾.

C'est partant d'un même constat d'absence de subventions publiques qu'Hubert Villard, directeur de la Bibliothèque cantonale et universitaire de Lausanne, annonce en 2007 l'accord passé entre l'État de Vaud et Google, portant sur la numérisation d'un lot de 100'000 volumes imprimés anciens du 17^e au 19^e siècle. Cela étant réalisé hors du pays, sans autres frais pour la collectivité que la mise à disposition de ressources humaines pour préparer et livrer les documents, ainsi que fournir les métadonnées nécessaires pour les signaler. Ce faisant, cette bibliothèque est devenue la première dans l'espace francophone à coopérer avec Google⁽³⁾.

Le paysage numérique est alors principalement constitué de serveurs institutionnels (*institutional repositories*) mis en place dans le sillage du mouvement Open access, à partir desquels les bibliothèques universitaires diffusent leurs publications et leurs thèses. Ce sont essentiellement des documents numériques natifs et non des documents numérisés à partir d'un original analogique. En 2004, le consortium Rero avait ouvert un serveur original faisant office à la fois de serveur institutionnel pour ses membres et de réservoir pouvant accueillir des documents patrimoniaux numérisés, amorce d'une véritable bibliothèque numérique⁽⁴⁾. Toutefois, celle-ci est restée modeste, proposant en été 2007 environ 5'000 documents.

À cela s'ajoute le *Catalogue collectif suisse des affiches*, une base iconographique qui compte alors quelque 40'000 notices accompagnées de reproductions numériques. Géré par la Bibliothèque nationale, ce serveur a reçu l'appui de la Fondation Memoriav.

La même année 2007 la Bibliothèque nationale, en collaboration avec Rero, met en place Digicoord, une plate-forme ouverte à toute institution afin de recenser les projets de numérisation existants. Digicoord offre ainsi un premier niveau de coordination, en réunissant et en partageant l'information existante sur tous les projets, qu'ils soient en cours de réalisation ou encore à l'état d'intention.

La Conférence universitaire suisse (CUS) lance alors le programme e-lib.ch. Le financement est assuré par la Confédération pour la période 2008-2011, au titre de soutien à la politique d'information scientifique et technique qui s'appuie sur les bibliothèques universitaires⁽⁵⁾. L'objectif général d'e-lib.ch est de mettre en place un portail national facilitant le travail de la recherche et l'accès à l'information.

À côté de projets portant sur les infrastructures ou visant à faciliter l'utilisation des ressources informationnelles (comme le métacatalogue Swiss-Bib ou le portail cartographique Kartenportal), plusieurs projets de numérisation sont validés:

- e-rara.ch: livres anciens
- e-codices.ch: manuscrits médiévaux
- retro.seals.ch: magazines et revues scientifiques suisses.

Après les programmes dont les sciences "dures" ont été largement bénéficiaires, notamment la coordination des ressources électroniques par le Consortium des bibliothèques universitaires suisses⁽⁶⁾, cet accent mis sur la numérisation constitue un rééquilibrage en faveur des sciences humaines et sociales. Les documents numérisés, matériaux de base pour les chercheurs, deviendraient non seulement facilement accessibles, mais ils y permettraient de nouvelles exploitations des contenus, notamment par le biais des "humanités digitales".

2. L'organisation du projet

Le projet e-rara.ch se donne pour mandat d'installer et de mettre en production un serveur, afin de présenter sur Internet des contenus numérisés de livres anciens en rapport avec la Suisse, de préciser ces contenus et de coordonner la numérisation des collections des bibliothèques.

Ces objectifs correspondent aux *Principes de Lund* établis en 2001 par la Commission européenne pour la numérisation: créer dans chaque pays un centre de compétences, s'appuyer sur les bonnes pratiques du domaine et coordonner les initiatives.

La Bibliothèque de l'Ecole polytechnique fédérale de Zurich (ETH) prend la responsabilité du projet, auquel sont associées les bibliothèques universitaires de Bâle et de Berne, la Bibliothèque de Genève, ainsi que la Zentralbibliothek de Zurich. En Suisse les bibliothèques universitaires abritent les collections patrimoniales de leur canton respectif^[7].

Les imprimés suisses du 16^e siècle représentent le corpus de départ, auquel la Bibliothèque de l'ETH ajoute une collection spécifique de textes fondamentaux en histoire des sciences. A côté de la Bibliothèque de l'ETH, les quatre institutions partenaires conservent la majeure partie de cette production imprimée. De façon significative, la Bibliothèque nationale suisse n'est pas associée, du fait de sa fondation relativement récente, en 1895.

Concrétiser cette entreprise suppose d'abord de répondre à de nombreuses questions ouvertes, sous-tendues par des enjeux aussi bien techniques qu'organisationnels, ce qui fait d'e-rara.ch un projet complexe et exigeant. C'est pourquoi les tâches à accomplir ont été réparties entre quatre groupes de travail, coordonnés par la Bibliothèque de l'ETH: contenu, métadonnées, numérisation, plate-forme. Chaque groupe comprend des représentants de chaque bibliothèque partenaire.

3. Contenu

Il est question de numériser un ensemble cohérent: un exemplaire de chaque titre imprimé en Suisse au 16^e siècle. À cette époque les villes ayant abrité des ateliers d'imprimerie sont peu nombreuses, et l'ensemble, évalué à 15'000 titres, est de taille raisonnable. Cette délimitation suit la pratique des grands travaux bibliographiques allemands sur les livres anciens, selon des tranches par siècles: 16^e siècle, puis 17^e et 18^e siècles (VD16, VD17, VD18)^[8].

Au cours du projet et des progrès des recensements, les quantités de livres imprimés au 16^e siècle dans les différentes villes évoluent. Par ordre décroissant:

Bâle	8'200 titres
Genève	3'250 titres
Zurich	1'626 titres
Berne	280 titres

Autres villes romandes:

Fribourg	66 titres
Lausanne	60 titres
Morges	41 titres
Neuchâtel	30 titres

Tableau 1. Evaluation du nombre de livres produits par les imprimeurs en Suisse au 16e siècle (Recensement du groupe de travail Contenu)

Les bibliothèques cantonales conservent les principales collections de livres imprimés sur leur territoire. Le recours aux autres bibliothèques partenaires permet de combler en partie certaines lacunes.

Les corpus allemands et suisses se recoupent car VD16 inclut toute la production dans l'espace géographique de langue allemande, dont les villes de Bâle, Berne, ou Zurich, mais également Fribourg ou Pruntrut (Porrentruy). Un contact avec la Bayerische Staatsbibliothek de Munich a très vite été établi pour assurer la coordination de la numérisation entre les deux projets et mutualiser ainsi les efforts.

La Bibliothèque de Genève a pris en charge la coordination de la numérisation de toutes les impressions romandes, en prenant contact avec les bibliothèques cantonales de Fribourg, de Lausanne et de Neuchâtel.

Malgré les synergies offertes par cette dynamique de réseau, une collection numérique exhaustive du 16e siècle suisse ne pourra pas être obtenue dans les limites temporelles définies au départ. Par exemple, près de 900 impressions genevoises de cette époque ne sont connues qu'au travers d'exemplaires dispersés dans un grand nombre de bibliothèques étrangères. Une tentative d'acquisition de reproductions numériques auprès de ces institutions s'est avérée rapidement très compliquée, en raison de diverses réticences, du nombre de démarches à entreprendre et du coût.

4. Métadonnées

La description des contenus numérisés par les métadonnées réunit deux types d'information: bibliographique (fiche signalétique des documents) et structurelle (représentant le contenu du document). Ces informations fournissent autant de points d'accès à la ressource, via le moteur de recherche interne au portail et, si elles peuvent être indexées, par un moteur de recherche généraliste sur le Web.

Les *métadonnées bibliographiques* sont extraites des catalogues de référence des bibliothèques partenaires, au travers du système Ex Libris Aleph, qui équipe les réseaux IDS Bâle-Berne et IDS Nebis⁽⁹⁾ ou VTLS Virtua, qui équipe le réseau Rero. S'y ajoute la base GLN15-16, base de données bibliographique exhaustive et spécialisée couvrant les

impressions des 15e et 16e siècles de la Suisse romande réformée, soit Genève, Lausanne, Neuchâtel, ainsi que Morges⁽¹⁰⁾.

La plate-forme de gestion numérique choisie, Visual Library de l'entreprise allemande Semantics, n'est pas conçue pour saisir directement les métadonnées bibliographiques. Un flux de données est établi avec les catalogues sources des bibliothèques au travers des protocoles d'échange automatique standards (Z39.50, OAI/PMH, SRW/SRU). Cela permet d'assurer que les notices descriptives dans e-rara.ch sont toujours à jour par rapport au catalogue de référence, dans lequel interviennent toutes les mutations. La base GLN15-16, issue d'un développement spécifique, ne permet pas l'échange automatique de données normalisées. Les informations bibliographiques doivent être importées manuellement à partir d'un fichier CSV, ainsi que les éventuelles mises à jour.

Contrairement aux métadonnées bibliographiques, les *métadonnées structurelles* ne sont pas préexistantes et doivent être saisies dans l'interface Visual Library. La structuration intervient après que les informations bibliographiques et l'ensemble des fichiers, sous forme de pages-images, ont été importés. Elle consiste à créer une table des matières, en regroupant les pages au sein de sections et d'éventuelles sous-sections préalablement repérées, facilitant ainsi l'usage de l'information. D'autres éléments potentiellement utiles sont également signalés, comme la page de titre ou divers éléments techniques.

Le résultat de la structuration peut être mis en forme selon des normes en vigueur, telles que METS⁽¹¹⁾ ou même TEI⁽¹²⁾, ce qui ouvre vers des exploitations variées: fourniture de capsules d'archivage ZIP (comprenant sous forme compressée les fichiers image accompagnés de leurs métadonnées), échange ou moissonnage des métadonnées par d'autres serveurs comme par des agrégateurs (par exemple le portail Gallica).

La structuration n'est pas une opération triviale pour des textes du 16e siècle: le traitement de chaque titre prend du temps et exige des connaissances suffisantes en langues anciennes et en compréhension de l'organisation du texte. C'est pourquoi elle n'est pas toujours appliquée, notamment pour les textes courts.

Un autre écueil tient à l'absence de texte intégral généré par OCR. Ce procédé, extrêmement précieux car il facilite la "recherchabilité" de l'œuvre dans un portail ou sur Internet ainsi que le repérage de séquences de mots à l'intérieur de l'œuvre, donne des résultats mitigés pour les textes anciens. La typographie et l'orthographe de l'époque, la qualité (contraste) variable de l'impression, le jaunissement et les déformations du papier, ont un impact négatif sur le taux d'exactitude.⁽¹³⁾ C'est pourquoi e-rara.ch a renoncé provisoirement à l'OCR. La question reste ouverte dans les plans de développement.

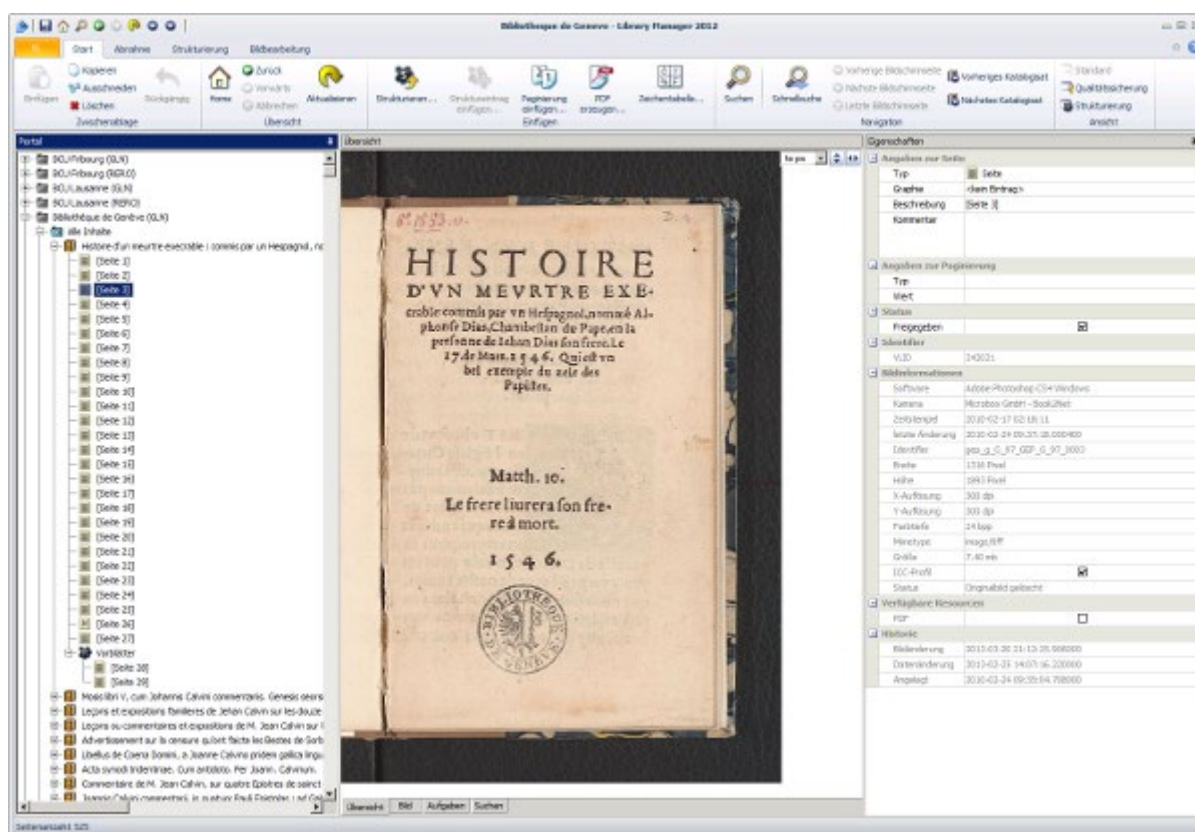


Illustration 2. Plate-forme e-rara.ch: interface professionnelle Visual Library Manager pour la structuration (<http://dx.doi.org/10.3931/e-rara-1036>)

5. Numérisation

Le groupe de travail Numérisation avait plusieurs missions: fixer les normes spécifiques pour les besoins du projet, évaluer puis acquérir deux types de scanners (manuel et automatique). Appuyé sur les bonnes pratiques, et notamment celles de la DFG⁽¹⁴⁾, le groupe considère que pour le livre ancien l'accès au contenu informationnel (texte, mais aussi schémas, équations...) ne suffit pas, la restitution de la matérialité de l'exemplaire revêt aussi une certaine importance. Les paramètres suivants sont donc adoptés:

- Résolution optique 300/400 dpi
- Numérisation en couleur
- Format d'enregistrement Tiff non compressé, profil ICC attaché
- Numérisation avec marge extérieure de 5 mm, afin de prouver l'intégrité du document
- Numérisation intégrale, y compris pages de garde, plats extérieur et intérieur de la reliure, dos, tranche.
- Ces normes sont complétées par des règles de nommage des fichiers, respectant l'ordre naturel de lecture⁽¹⁵⁾, et un contrôle de la qualité.

Outre le respect des paramètres de numérisation susmentionnés, l'évaluation des scanners a porté particulièrement sur les critères suivants: préservation des livres anciens, qualité

colorimétrique, netteté, rapidité du processus de numérisation, ergonomie. Si le scanner manuel choisi et qui a équipé chaque bibliothèque partenaire (cf. illustration 3) a globalement répondu aux attentes, les deux scanners automatiques installés dans les ateliers de Bâle et Zurich ont déçu en termes de productivité. Avec les livres anciens, dont la manipulation est délicate, le mécanisme qui tourne les pages automatiquement n'offre pas une fiabilité suffisante et implique l'intervention permanente d'un opérateur. La ZB Zurich a ainsi estimé que sa productivité réelle sur cet appareil n'était que de 2'000 scans par jour, le quart de celle annoncée par le fabricant⁽¹⁶⁾.



Illustration 3. Numérisation à la Bibliothèque de Genève (photo M. Thomann, juin 2010)

Le projet e-rara.ch s'est développé durant une période où les techniques de numérisation ont considérablement évolué. D'un côté la popularité croissante des opérations de numérisation a élargi et stimulé le marché des scanners. En parallèle, les campagnes de numérisation dans les institutions se sont professionnalisées. Une prise de conscience accrue de l'importance du patrimoine a conduit à de plus grandes exigences quant aux mesures destinées à sa préservation.

L'intervention des professionnels de la conservation s'est généralisée. Ils interviennent en amont, examinant les documents à numériser et vérifiant que les contraintes mécaniques des scanners ne leur sont pas dommageables⁽¹⁷⁾. Les fabricants pour leur part développent de nouveaux appareils qui visent à simplifier et à accélérer les actions de l'opérateur, tout en préservant au maximum l'intégrité matérielle des documents. La principale difficulté posée par les livres anciens est liée à l'angle d'ouverture que permet l'état du volume. Plus l'angle est réduit, moins nocive sera la numérisation. Zurich a acquis sur son propre budget un scanner nécessitant une ouverture de 30° seulement, rendant ainsi techniquement possible la numérisation de centaines de volumes supplémentaires (Illustration 4).



Illustration 4. Capter l'information dans les livres anciens exige des scanners spécialisés (Anagramm 30°)

Cette évolution de la numérisation et de la valeur attachée au patrimoine a eu des répercussions sur l'organisation générale du projet. Prévu initialement autour d'un atelier central situé à Zurich, qui aurait traité la majorité des collections des autres partenaires, le schéma a évolué vers une production plus décentralisée. Cela en raison de la complexité croissante (administrative, liée aux mesures de sécurité) du transport de livres pour la plupart rares et précieux, sans parler d'une certaine réticence des responsables au regard des risques encourus.

6. Plate-forme

Le dernier groupe devait mettre en place un système complet de "bibliothèque numérique", hébergé physiquement au centre informatique de la Bibliothèque de l'ETH. Par système complet il faut comprendre une plate-forme capable de gérer les objets numériques et de les présenter au public sur le Web. Les fonctionnalités attendues s'appliquent à ces deux volets. Pour les professionnels, cela concerne la saisie des métadonnées ainsi que le classement en collections. L'interface publique doit être orientée utilisateurs en termes de recherche, de navigation et de récupération des fichiers. L'attribution d'un identifiant à chaque document numérique assure la pérennité de l'accès à la ressource tout en permettant de la citer commodément. Le système DOI a été choisi à cet effet⁽¹⁸⁾. Une interface d'échange de données standardisées permet d'intégrer le contenu d'e-rara.ch dans d'autres portails ou agrégateurs.

Le système Visual Library, développé par les entreprises Semantics et Walter Nagel, équipe plusieurs bibliothèques numériques, principalement en Allemagne. Son principe d'organisation des documents par "domaines" autonomes convenait bien aux besoins de gestion décentralisée par bibliothèques.

Un problème récurrent dans la conception des bibliothèques numériques est la dérivation, à partir des images master au format Tiff, des images numériques de moindre résolution adaptées à l'affichage Web. Cela est entièrement pris en charge par Visual Library après l'importation des fichiers (par transfert FTP ou envoi de disques durs), ainsi que la génération des fichiers PDF (totalité du document ou par chapitre), sans intervention particulière de l'opérateur du moment que celui-ci active le bouton "publication".

Après importation et publication, le système crée sur demande une "capsule ZIP", soit la réunion dans un dossier compressé d'un ensemble de fichiers images Tiff accompagnés des métadonnées issues du catalogue et de la structuration. Ces capsules ZIP sont des entités pouvant rejoindre des serveurs d'archivage numérique à long terme, tâche qui est de la responsabilité de la bibliothèque possédant le document d'origine.

L'illustration 5 montre le workflow global des opérations. Le tableau qui suit synthétise le nombre de titres et de pages intégrés chaque année. Le nombre de pages par livre varie mais représente une moyenne de 300.

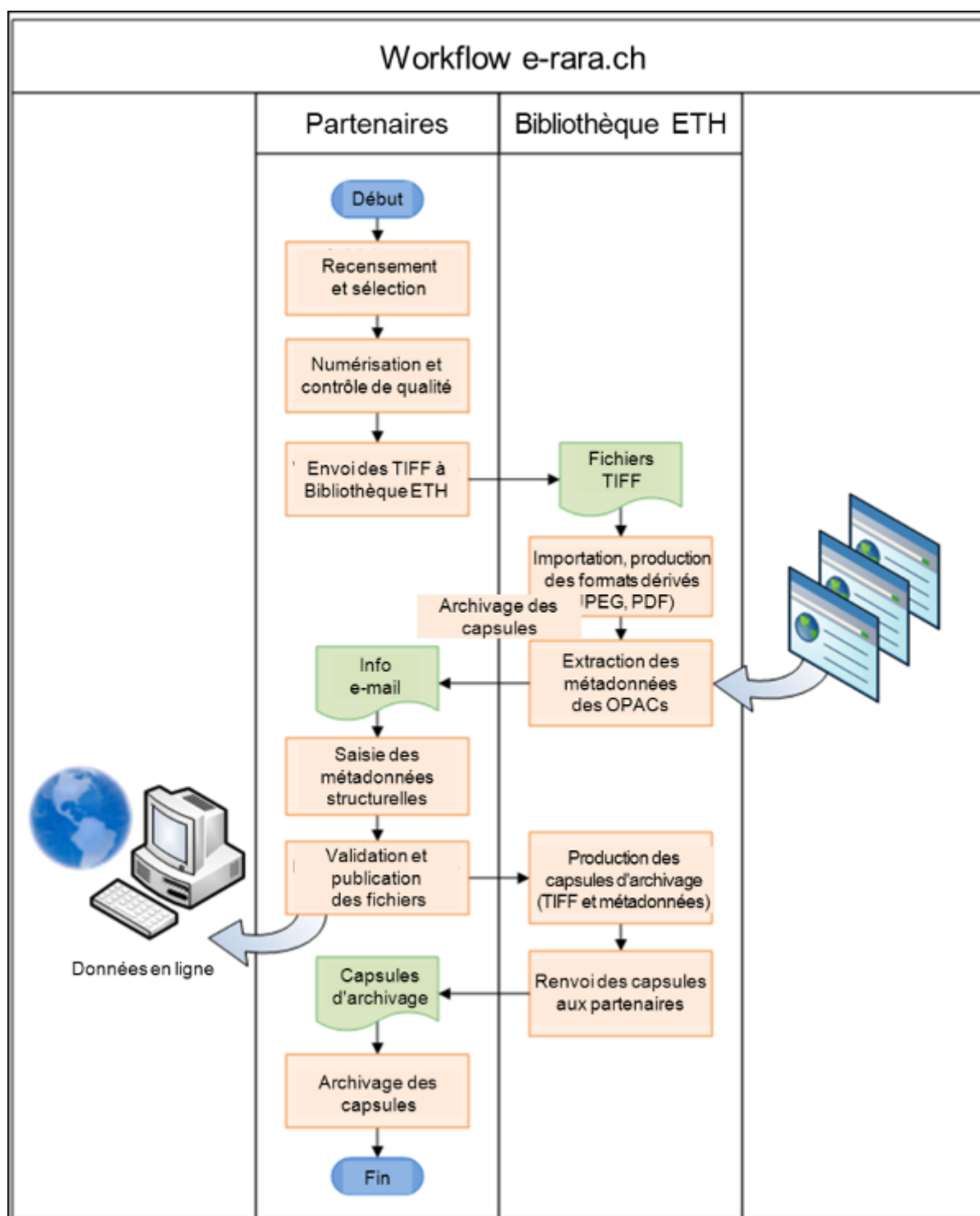


Illustration 5. Workflow général du projet e-rara.ch.

2008	—	—	—
2009	1'686 titres	504'911 pages	299 pages par titre
2010	3'496 titres	1'194'0905 pages	342 pages par titre
2011	3'323 titres	1'128'805 pages	340 pages par titre
2012	6'854 titres	2'072'486 pages	302 pages par titre
2013	7'156 titres	1'660'866 pages	232 pages par titre
2014 (jusqu'au 10.11.2014)	6'260 titres	1'955'137 pages	312 pages par titre

Tableau 2. Production pour e-rara.ch des bibliothèques partenaires en 2008-2014
(Source: Coordination e-rara.ch).

7. MISE EN SERVICE DU PORTAIL

E-rara.ch ouvre au public début 2010, deux ans après le lancement du projet, avec une offre initiale de 500 titres. L'intérêt suscité dès les premiers jours par cette nouvelle bibliothèque numérique se confirme par la suite, avec un nombre mensuel de visites en croissance régulière, à mesure que l'offre s'étoffe (Illustration 6).

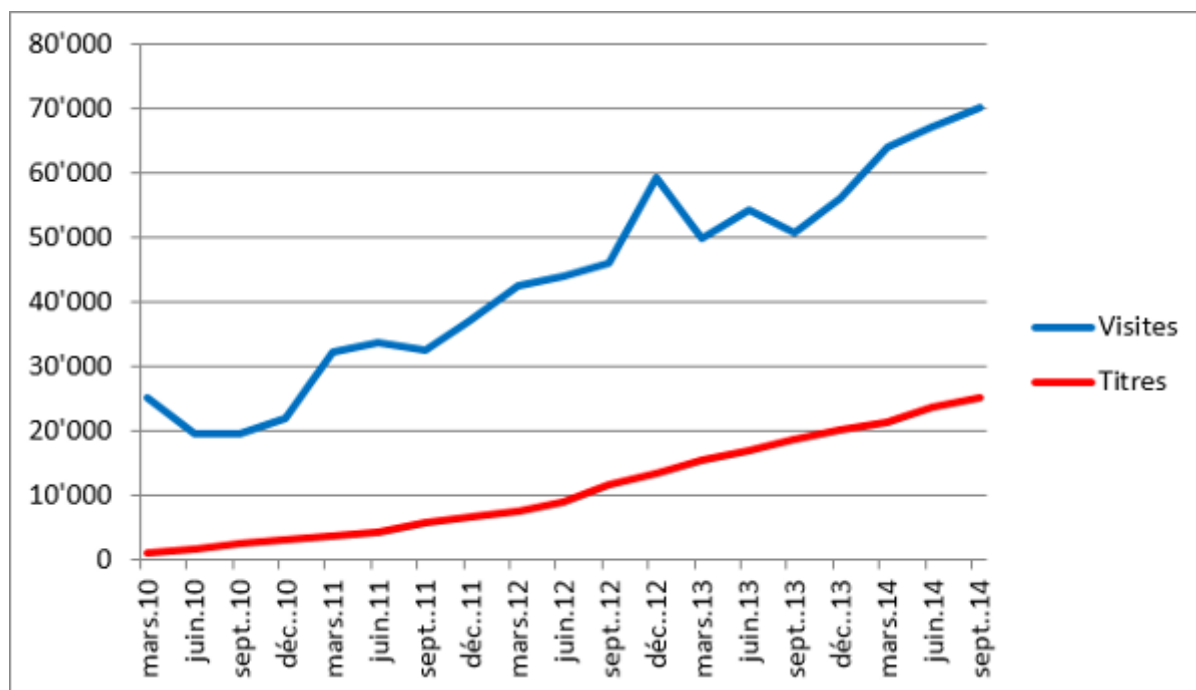


Illustration 6. Fréquentation mensuelle du portail et nombre de titres disponibles (vue par trimestre).

Outre un bon référencement dans les moteurs de recherche, l'offre est également visible dans les métacatalogues, comme Swissbib⁽¹⁹⁾, et les outils de découverte de ressources des partenaires comme Wissensportal (à Zurich)⁽²⁰⁾ ou Explore (Rero)⁽²¹⁾. En outre le serveur est régulièrement moissonné par Gallica, la bibliothèque numérique française de référence⁽²²⁾, ainsi que par la base Eromm⁽²³⁾. Des liens réciproques vers le répertoire VD-16 pour le domaine germanique ou GLN 15-16 pour la Suisse romande sont également en place.

Les données d'accès par pays montrent clairement le rayonnement international d'e-rara.ch (illustration 7). La première place occupée par l'Allemagne n'est pas étonnante, étant donné la proportion élevée de documents provenant des bibliothèques alémaniques. L'Allemagne, les États-Unis, la Suisse et la France représentent presque 60% de la fréquentation totale du site.

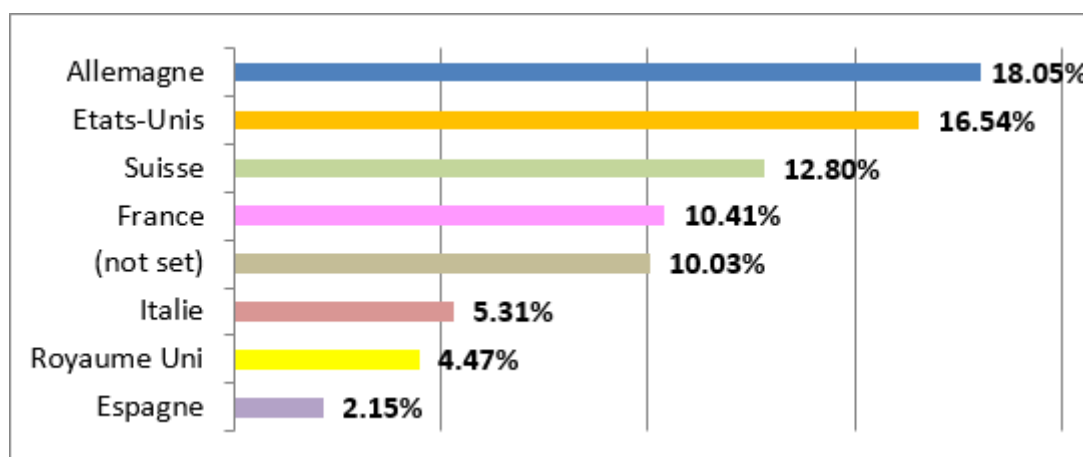


Illustration 7. Origine des internautes, janvier-septembre 2014

8. Évolution du projet

Contrairement à une publication imprimée qui représente l'état d'une recherche à un moment donné, un serveur Internet qui n'évoluerait pas perdrait assez rapidement son audience. Dès le début du projet l'évolution de la plateforme e-rara.ch a été envisagée, tant vis-à-vis du contenu et de son étendue que vis-à-vis du contenant, c'est-à-dire les fonctionnalités et l'ergonomie de l'interface utilisateur.

La définition des contenus éligibles pour figurer dans e-rara.ch s'est successivement élargie sur plusieurs plans:

- La période couverte

Initialement limitée au 16^e siècle, elle englobe maintenant ce que les bibliothèques désignent comme livres anciens⁽²⁴⁾, soit les impressions du 15^e au 19^e siècle. Des livres du début du 20^e siècle pourraient aussi être accueillis pour autant qu'ils soient dans le domaine public.

- Les types de documents

De nouvelles collections ont été créées pour intégrer, à côté des livres, les cartes et les partitions musicales. Si la question de traiter des périodiques ne se pose pas (d'autres plateformes spécifiques existent⁽²⁵⁾), inclure des estampes ("Druckgraphik") est en cours de réflexion. Le périmètre défini est celui des documents imprimés au sens large, par opposition aux manuscrits ou documents d'archives.

- Les impressions hors de Suisse

La Bibliothèque de l'ETH possède un vaste fonds patrimonial scientifique, mais peu de livres anciens imprimés en Suisse. Elle a donc contribué à la numérisation de documents édités hors de Suisse, réunis dans la collection "Textes pour l'histoire des sciences". D'autres ensembles ont suivi, comme "Poésie et prose italienne du 17^e siècle" numérisée par la Biblioteca Salita dei Frati à Lugano. La seule limite fixée est que les livres doivent se trouver dans une bibliothèque suisse.

- Les contributeurs

E-rara.ch ayant élargi le périmètre des documents proposés est devenue en peu d'années une ressource incontournable au plan national. La plate-forme se prépare donc à intégrer d'autres bibliothèques qui souhaiteraient y participer.

Pour ce qui est du contenant, une étude d'utilisabilité – c'est-à-dire l'évaluation de l'efficacité des recherches et de la satisfaction de l'utilisateur – a été conduite en 2011⁽²⁶⁾. Elle a donné lieu à une refonte importante de l'interface web, mise en ligne en 2013. Grâce aux filtres par facettes, dispositif qui s'est généralisé ces dernières années, il est possible de préciser la recherche lorsque les résultats sont très nombreux. La zone de recherche elle-même est mieux mise en évidence. Un point important concerne l'abandon sur le bandeau des rubriques liées aux institutions partenaires. Elles permettaient d'afficher la liste de tous les livres conservés par une de ces bibliothèques, ce qui est compréhensible du point de vue de l'institution en quête de visibilité sur le Web, mais n'a absolument aucun sens pour l'utilisateur cherchant des contenus.

Le dynamisme technologique impose un rythme soutenu et oblige les systèmes à s'y adapter, sous peine de disparition rapide. Sur ce plan, e-rara.ch devra impérativement améliorer son interface sur les dispositifs mobiles comme les tablettes. L'absence de texte intégral généré par OCR, évoqué plus haut, sera aussi difficilement défendable à long terme.

Le manque de renvois entre les différentes formes des noms propres (Aristote en français, Aristotle en anglais, Aristoteles en allemand et parfois aussi en anglais, Aristotele en italien, etc.) produit du silence dans les réponses et complique la mise en oeuvre des listes alphabétiques ou des facettes. L'implantation d'un thésaurus établissant des équivalences a une importance particulière pour un site dédié aux imprimés anciens et devrait être réévalué⁽²⁷⁾. A moyen terme, on peut imaginer que les techniques du *Web de données*⁽²⁸⁾ apporteront des solutions plus légères pour intégrer ces disparités linguistiques.

9. Bilan et perspectives pour l'avenir

En 2009, dans un dossier traitant des politiques de numérisation, un éditorialiste soutient que la firme "Google n'accapare rien, elle profite du vide public"⁽²⁹⁾. La réalisation d'e-rara.ch est le sursaut né d'une prise de conscience politique, prouvant qu'il est possible d'occuper le terrain lorsque des moyens suffisants y sont consacrés.

La bibliothèque numérique e-rara.ch est le fruit, comme dans bien d'autres domaines, d'une utilisation pragmatique des compétences et des moyens existant dans le pays. Bien que la Bibliothèque nationale ait la mission de coordonner les initiatives liées aux bibliothèques, c'est la Bibliothèque de l'ETH qui a porté le projet d'un serveur national pour les livres anciens numérisés, dans le cadre d' e-lib.ch. Forte de plusieurs réalisations antérieures dans le domaine numérique (Consortium des bibliothèques universitaires suisses, numérisation des périodiques scientifiques), elle a pu y renforcer son rôle de leader.

Puisque le recul le permet, il est bon de rappeler les facteurs de cette réussite, même si certains d'entre eux peuvent paraître évidents:

1. La méthode projet s'est révélée très efficace. Elle a bénéficié d'un apport financier fédéral bienvenu, les partenaires mettant à disposition essentiellement des ressources en personnel. Les synergies ont été profitables à de nombreux points de vue: apport d'une

masse critique pertinente de livres anciens, mise en commun des expertises, économies d'échelle pour la commande d'équipements (scanners, écrans), rigueur des évaluations.

2. Une certaine faculté d'adaptation et de redistribution des moyens a permis un grand gain de temps et une bonne cohérence dans la planification du projet. Par exemple, des équipements supplémentaires ont pu être prêtés par Zurich pour être mis à disposition à Genève, en raison de l'ampleur des collections à numériser.
3. Les exigences de qualité étaient élevées, mais raisonnablement atteignables. L'adaptation aux normes usuelles a permis de connecter immédiatement le portail à d'autres réseaux de ressources.
4. Le résultat est un portail d'envergure nationale, réunissant un vaste ensemble de documents présentés de façon identique, quelle que soit la bibliothèque de provenance.

Trop souvent en effet les institutions, par peur d'une dissolution de leur identité, tiennent à disposer de leur propre bibliothèque numérique pour mettre en valeur les richesses de leur ville ou de leur région. Cela n'apporte en fait que peu de bénéfices. Raphaële Mouren le souligne ainsi: "Les bibliothécaires savent pourtant que les portails des bibliothèques sont très peu utilisés par les internautes, et que les chercheurs, comme les autres publics, ne s'amuse pas à consulter l'un après l'autre les sites des bibliothèques pour y chercher leur bonheur: comme tout le monde, ils utilisent les moteurs de recherche."⁽³⁰⁾

Denis Roegel, un autre chercheur grand connaisseur des bibliothèques numériques, fustige aussi la disparité des sites institutionnels⁽³¹⁾. En outre, il met en relief un manque de réflexion préalable, rendant parfois peu pertinente l'utilisation des ressources. Une meilleure coordination entre bibliothèques, notamment pour ce qui est des critères de qualité requis, éviterait de mettre en ligne, comme c'est encore trop souvent le cas, des numérisations insuffisantes pour les chercheurs et donc finalement peu utiles.

Dans le sillage d'e-rara.ch et utilisant le même système, une nouvelle plate-forme est créée pour les manuscrits et archives: e-manuscripta.ch. Cette infrastructure, également nationale, est employée par les bibliothèques universitaires de Bâle et de Berne. L'apport principal de contenus viendra cependant de la ZB Zurich, qui avec son projet DigiTUR collabore étroitement avec l'ETH. En 2012 un financement important (près de 10 millions de francs) lui a été octroyé par le Fonds de loterie⁽³²⁾ (*Lotteriefonds*) pour la numérisation de ses collections patrimoniales⁽³³⁾. On peut ainsi dire que le pôle de Zurich gère une infrastructure à vocation nationale tout en répondant largement à ses propres besoins. A l'instar de Zurich, Genève, deuxième ville du pays et centre historique du livre, développe une stratégie pour intensifier la présence de son patrimoine en ligne. Les autres régions (Berne, Vaud, Valais par ex.) s'y investissent également⁽³⁴⁾.

Le programme national qui a succédé à e-lib.ch pour la période 2013-2016 est "Information scientifique: accès, traitement et sauvegarde" (CUS-P2)⁽³⁵⁾. Il prend en compte le fait que la numérisation des bibliothèques nécessite un appui financier qui s'ajoute au financement propre des institutions. Toutefois les demandes déposées dans ce cadre concernant la numérisation ont toutes été suspendues, y compris la poursuite du développement d'e-rara.ch. La direction du programme désire en effet regrouper ces requêtes sous l'autorité d'un même comité de pilotage. Souhaitons que cela permette de véritables synergies nationales, tout en gardant comme visée le bien commun: que les solutions mises en place correspondent aux besoins de toutes les institutions et répondent aux attentes des utilisateurs.

10. E-rara.ch – dates clés

Août 2007

Demande de projet individuel dans le cadre du programme "E-lib.ch: Bibliothèque électronique suisse". Titre du projet: "www.e-rara.ch: die Online-Plattform für digitalisierte Rara aus Schweizer Bibliotheken" (www.e-rara.ch: portail des livres anciens numérisés pour les bibliothèques suisses).

Novembre 2007

Acceptation du projet e-rara.ch par la Conférence universitaire suisse (CUS).

Janvier 2008

Lancement du projet.

Octobre 2009

Mise en production de la plate-forme pour les partenaires (système Visual Library de l'entreprise Semantics).

15 mars 2010

Ouverture officielle du portail internet au public.

2012

Première prolongation du projet (financement complémentaire).

Été 2012

Le seuil de 10'000 documents consultables en ligne est franchi.

2013

Seconde prolongation du projet (financement complémentaire).

Janvier 2013

Evolution majeure de l'interface: champ de recherche et navigation par facettes.

2014

Première année d'exploitation de la plate-forme sans financement CUS.

Mars 2014

Demande de projet dans le cadre du nouveau programme de la CUS pour les années 2013-2016 (CUS-P2): "Information scientifique: accès, traitement et sauvegarde".

Novembre 2014

Evolution légère de l'interface. Le portail offre plus de 26'000 documents à la consultation.

Dès 2015

Adaptation de l'interface pour les appareils mobiles, intégration du texte intégral généré par OCR, élargissement des institutions participantes

NOTES

- (1) Sur le projet e-rara.ch, voir: Franziska Geisser, "E-rara.ch: ein Schweizer Digitalisierungsprojekt mit internationaler Ausstrahlung", in *Arbido*, (3) 2011, p. 23-26. Trad. fr. par Alexis Rivier, *E-rara.ch: un projet suisse de numérisation au rayonnement international*. <http://doc.rero.ch/record/27345>
- (2) Doffey Marie-Christine, in *Le Temps*, 4.5.2007.
- (3) "La Bibliothèque cantonale et universitaire vaudoise rejoint le projet Recherche des Livres de Google", Bureau d'Information et de Communication de l'Etat de Vaud, 15.5.2007.
- (4) Rero Doc, <http://doc.rero.ch>, bibliothèque numérique du Réseau des bibliothèques de Suisse occidentale (Rero), <https://www.rero.ch>.
- (5) Mise en place par la *Conférence universitaire suisse* d'une *Commission pour les bibliothèques universitaires* (CBU) en 1983 et d'une Commission pour l'information scientifique dès 1985. Voir notre article, "Vingt ans de nouvelles technologies dans les bibliothèques suisses: des références aux contenus numériques", in *Arbido print*, (2) 2007, p. 26-34.
- (6) Coordonne la fourniture de ressources électroniques depuis 2000. <http://lib.consortium.ch>.
- (7) Le cas est similaire en Allemagne: les bibliothèques universitaires conservent souvent le patrimoine de Länder.
- (8) VD16: Verzeichnis der im deutschen Sprachbereich erschienenen Drucke des 16. Jahrhunderts (Répertoire des livres imprimés dans les pays de langue allemande au 16e siècle). Depuis 2006, la DFG (Deutsche Forschungsgemeinschaft - l'Agence allemande pour la recherche scientifique) finance la numérisation des titres recensés dans VD16. La Bayerische Staatsbibliothek de Munich, dont les collections de livres anciens sont parmi les plus étendues, est leader de ce projet. www.bsb-muenchen.de/index.php?id=1681.
- Voir aussi: Sommer Dorothea, "Digitalisierung von Drucken des 17. Jahrhunderts an der Universitäts- und Landesbibliothek Halle: ein Werkstattbericht zu einem DFG-Projekt der Aktionslinie VD 16/ VD 17", in *ABI-Technik*, 27 (4), 2007, p. 236-247.
- (9) IDS: Informationsverbund Deutschschweiz. <http://www.informationsverbund.ch>.
Nebis: Netzwerk von Bibliotheken und Informationsstellen in der Schweiz. <http://www.nebis.ch>.
- (10) Cette base est l'œuvre du bibliographe belge Jean-François Gilmont, spécialiste de l'histoire du livre au temps de la Réforme en Suisse romande. <http://www.ville-ge.ch/bge/gln>.
- (11) Metadata Encoding and Transmission Standard. Norme gérée par la Library of Congress pour décrire dans le formalisme XML des objets numériques complexes (livres). <http://www.loc.gov/standards/mets>.
- (12) Text Encoding and Interchange. Norme d'encodage pour les textes littéraires et l'édition scientifique. <http://www.tei-c.org>.
- (13) On estime qu'un taux d'exactitude de 90-95% est nécessaire pour la recherche. Plusieurs projets se consacrent à l'amélioration de la qualité de l'OCR. Par ex. IMPACT (IMProving ACcess to Text), <http://www.impact-project.eu> et SUCCEED (Support Action Centre of Competence in Digitisation), <http://www.succeed-project.eu>.
eMOP (Early Modern OCR Project), <http://idhmc.tamu.edu/emop>.

- (14) DFG-Praxisregeln "Digitalisierung", Deutsche Forschungsgemeinschaft, 2013. Également en version anglaise: *DFG Practical Guidelines on Digitisation*. http://www.dfg.de/formulare/12_151.
- (15) Pour les imprimés en hébreu notamment, la numérisation respecte l'inversion du sens de lecture. Par ex. <http://dx.doi.org/10.3931/e-rara-3068>.
- (16) Edgar Schuler, "Die ZB ist Google für alte Schweizer Drucke", in *Tages Anzeiger*, 30.7.2009.
- (17) Les implications de la numérisation en terme de conservation sont bien décrits dans: Anna E. Bulow et Jess Ashmon, *Preparing Collections for Digitization*, London, Facet Publishing, 2011.
- (18) La Bibliothèque de l'ETH est agence d'attribution pour la Suisse des DOI (Digital Object Identifier). <http://www.doi.org>.
- DOI Desk: <https://www.library.ethz.ch/de/Dienstleistungen/Publizieren-registrieren-verwalten/DOI-Desk-der-ETH-Zürich>.
- (19) <https://www.swissbib.ch>.
- (20) <http://www.library.ethz.ch>.
- (21) <http://explore.rero.ch>.
- (22) <http://gallica.bnf.fr>. Les statistiques d'accès, montrent que chaque année plusieurs milliers de visites de e-rara.ch proviennent d'une recherche dans Gallica.
- (23) European Register of Microform and Digital Masters. Fondée en 1994 comme catalogue collectif européen pour les copies de substitution sur microfilms, elle recense également depuis les œuvres numérisées. <http://www.eromm.org>.
- (24) La date limite fluctue d'une institution à l'autre: 1850 ou 1900 sont usuelles.
- (25) retro.seals.ch déjà mentionné plus haut pour les revues numérisées et Presse suisse en ligne pour les journaux. <http://newspaper.archives.rero.ch>.
- (26) Eliane Blumer, *Benutzerorientierte Evaluation der Webseite e-rara.ch anhand von Usability- und Blickmessungstests*, Mémoire de bachelor, Genève, Haute école de gestion, 2011. <http://doc.rero.ch/record/28011>.
- (27) Notamment mise en évidence dans une évaluation plus récente du portail e-rara.ch: Meystre Valérie, *Tests d'utilisabilité: comparaison de deux méthodes appliquées au site e-rara.ch*, Travail de recherche (Master), Haute école de gestion de Genève, 2014. <http://doc.rero.ch/record/209599>.
- (28) En anglais *Linked data*, à rapprocher également du web sémantique. <http://linkeddata.org>.
- (29) Nicolas Dufour, "Le grand vide public", in *Le Temps*, 19.9.2009.
- (30) Raphaële Mouren, "Les manuscrits médiévaux dans les bibliothèques numériques en France", in Fabienne Henryot (éd.), *L'historien face au manuscrit: du parchemin à la bibliothèque numérique*, Louvain-la-Neuve, Presses universitaires de Louvain, 2013, p. 343-355.
- (31) Denis Roegel, *La numérisation durable*, 21.12.2013. <http://locomat.loria.fr/other/2012numerisation-durable.pdf>.
- (32) Le *Lotteriefonds* est l'organe de répartition des bénéfices de loterie pour les cantons de Suisse alémanique et le Tessin.

⁽³³⁾ DigiTUR. <http://www.zb.uzh.ch/spezielsammlungen/digitur/index.html.de>.

Communiqué de presse, canton de Zurich, 28.6.2012, *Zentralbibliothek digitalisiert Zürcher Bestände*, <http://www.zh.ch/internet/finanzdirektion/de/aktuell.newsextern.-interne...>.

⁽³⁴⁾ Berne: DigiBern. <http://www.digibern.ch> Vaud: Scriptorium. <http://scriptorium.bcu-lausanne.ch> Valais: Vallesiana. <http://www.vallesiana.ch>.

⁽³⁵⁾ <http://www.crus.ch/information-programmes/projets-programmes/isci.html?L=1>.