



Virtuel Mucem

Remédiation des sites web de la collection
Recherches ethnologiques (2005-2008)

Alexandre Faye, Sara Aubry, Géraldine Camile, Sarah
Dietz, Charles Riondet, Chloé Sussan-Molson, Jasmine
Fejzovski

Publié le 15-06-2026

<http://sens-public.org/articles/1794>



Abstract

The Virtuel Mucem project (Mucem, BnF, AMu) addresses the preservation and remediation of digital heritage content produced in the 2000s and now lost due to technological obsolescence. This corpus of 8 websites, reflecting the Mucem's emerging identity (prior to its 2013 opening), was examined from multiple perspectives. Given their partial and fragmented archiving (BnF's WARC files, Mucem's "digital bulk"), the project combined an archival methodology (source inventory and cataloging, interviews with website producers), technical remediation (creation of a derived WARCD format, content indexing, secure replay environment), and editorialization (an interface for contextualizing the archives). (AF *et al.*)

Résumé

Le projet Virtuel Mucem (Mucem, BnF, AMu) aborde la préservation et la remédiation de contenus numériques patrimoniaux produits dans les années 2000 et disparus du Web, victimes de l'obsolescence technologique. Un corpus de 8 sites, reflétant l'identité naissante du Mucem (avant son ouverture en 2013) a été traité sous plusieurs angles. Face à leur archivage partiel et fragmenté (archives du web sous la forme de fichiers WARC à la BnF, « vrac numérique » au Mucem), le projet a combiné une méthodologie archivistique (recensement et inventaire des sources, entretiens avec les producteurs des sites), une remédiation technique (création de fichiers WARC dérivés, indexation des contenus, environnement sécurisé pour la rejouabilité) et une éditorialisation (interface de contextualisation des archives). (AF *et al.*)

Mot-clés : BnF, MuCEM, remédiation, préservation numérique, médiation culturelle

Keywords: BnF, MuCEM, remediation, digital preservation, cultural mediation

Table des matières

| | |
|---|----|
| Introduction | 4 |
| L'origine du projet | 4 |
| La collection <i>Recherches ethnologiques</i> | 6 |
| Contexte institutionnel | 6 |
| Une expérience de préfiguration virtuelle? | 8 |
| Traitement archivistique du corpus | 9 |
| Le dispositif technique de remédiation de l'archive | 11 |
| Produire une nouvelle archive web à des fins de remédiation à partir des collections de la BnF | 11 |
| Conception de l'outil d'extraction et création d'un environnement de consulta- tion sécurisé | 13 |
| Éditorialisation | 15 |
| Présentation de deux sites de la collection | 16 |
| Une journée de restitution pour redécouvrir les sites dans leur contexte | 18 |
| Pistes pour une reproductibilité | 21 |
| Conclusion | 22 |
| Bibliographie | 22 |

Virtuel Mucem

Alexandre Faye Sara Aubry Géraldine Camile Sarah Dietz
Charles Riondet Chloé Sussan-Molson Jasmine Fejzovski

Introduction

On sait que les bibliothèques brûlent, les musées aussi. La disparition de la documentation numérique, tout aussi réelle, est moins perceptible ; elle a longtemps été un impensé, sinon pour les archivistes, du moins pour une partie de ses producteurs. Les musées, chargés de conserver des collections physiques, sont aussi producteurs de recherche et de savoirs, qu'ils diffusent sous diverses formes. La plus évidente en est l'exposition, pensée et perçue comme temporaire ou évolutive, et son prolongement édité, durable, qu'en sont les catalogues. Depuis les années 1980 apparaissent de nouveaux formats numériques, en ligne ou non : bases de données telles que Joconde, publiée sur Minitel via *3614 Joconde* à partir de 1992, sites internet et expositions virtuelles dans les années 1990, CD-ROM et vidéodisques. Ces nouveaux documents échappent aux habitudes précédentes, le contenu n'étant accessible que via un dispositif technique (Pédaque 2006). Contrairement aux catalogues papiers, les sites internet n'ont pas de pérennité en eux-mêmes : leur contenu ne reste accessible qu'aussi longtemps que l'institution productrice maintient serveurs, noms de domaines, et technologies. Pour autant, le public peut s'en être emparé, sans que l'institution n'en ait toujours conscience. Le jour où l'accès au site n'est plus possible se pose la question de restituer, sous une forme ou une autre, ces ressources documentaires. L'archivage du web en est une solution privilégiée, mais ne répond pas à tous les besoins ni à toutes les questions. Comment garder la lisibilité d'un site développé avec Flash ? Pourquoi redonner accès à un site dont le contenu paraît daté ? Comment compléter les creux de ces archives ? Le Mucem a été amené à se poser ces questions lorsque des chercheurs ont fait part de leur déception à la mise hors ligne fin 2020 d'un site sur les cornemuses, produit en 2007 par l'ethnomusicologue Marie-Barbara Le Gonidec. Cette interpellation a été le point de départ d'une réflexion plus large, portée dans le cadre d'un appel à projet BnF Data Lab. Du fait du contexte singulier de production de ce site, la question technique du départ s'est élargie et déplacée, conduisant à retisser une histoire plus large, touchant à celle de l'institution comme des pratiques de médiation numérique des années 2000.

L'origine du projet

Très rapidement, il apparaît que le site *Cornemuses d'Europe et de Méditerranée* est intégré à une collection, les *Recherches ethnologiques*, produite durant les années de préfiguration du

Mucem¹. Les archives numériques de cette période sont parcellaires et peu structurées : les sites en eux-mêmes n’avaient pas été archivés, et la documentation du projet dispersée dans un « vrac numérique ». Pour répondre aux attentes de Marie-Barbara Le Gonidec, le Mucem s’est donc tourné vers d’autres partenaires. L’idée du projet Virtuel Mucem prit forme lors d’une discussion en marge de la 5^e conférence RESAW (Gebeil et Peyssard 2024) accueillie par le MucemLab en juin 2023. Puisque les sites avaient été archivés par la Bibliothèque nationale de France (BnF) dans le cadre de sa mission de dépôt légal de l’internet français², il devait être possible de fournir une reproduction au Centre de conservation et de ressources du Mucem. L’objectif principal était d’expérimenter une forme de remédiation des sites archivés en imaginant un dispositif éditorial et technique. Le regard porté sur ces matériaux passés est également au cœur du projet, car ces sites sont les composants d’une triple histoire : celle des enquêtes-collectes et des collections du musée, celle de l’internet et des médiations numériques, et celle d’une institution en devenir qui cherche à innover. Ainsi, notre projet ne se borne pas aux murs du Mucem, mais vise à positionner ces sites web comme des objets-témoins : on les considère comme représentatifs d’une certaine modalité d’investissement muséal et scientifique du web. Parce qu’ils portent la marque des circonstances de leur conception, il est nécessaire de les archiver adéquatement en tenant compte de leurs spécificités et de souligner leur historicité à travers le dispositif de remédiation.

Les objectifs du projet étaient à la fois documentaires et techniques. Il a été nécessaire de procéder à un traitement archivistique des sources locales, de les compléter par des enquêtes orales menées auprès des producteurs des sites et enfin d’intégrer les archives web de la BnF. En parallèle, une instruction technique a permis de réaliser un travail inédit d’empaquetage des archives du web afin de générer une archive de chacun des sites, qui soit bien délimitée et la plus complète possible. Suite à ces deux tâches menées simultanément, l’équipe projet a mis en place une interface éditorialisée de remédiation des sites pour consultation dans les murs du Centre de conservation et de ressources du Mucem.

Ce projet de remédiation a bénéficié du soutien du BnF Datalab et a été porté par trois partenaires : le Mucem, l’Université d’Aix-Marseille et la BnF³. Il porte une dimension collégiale et expérimentale. Il a rassemblé durant une année plus de quinze personnes,

1. Le site du Ministère de la Culture proposait trois collections en ligne : *Grands sites archéologiques*, *Célébrations nationales* et *Recherches ethnologiques*. Ces productions ont été conçues par des experts du domaine en lien avec l’équipe multimédia de la Mission de la recherche et de la technologie du Ministère de la Culture (MRT). Les huit sites qui constituent le corpus d’étude du projet Virtuel Mucem sont : *Cornemuses d’Europe et de Méditerranée – Les voyages du verre – Café, Cafés – Féminin / Masculin – Trésors du quotidien Europe et Méditerranée – Loïn de l’Ararat – Hip-Hop, art de rue, art de scène – L’olivier, trésor de la Méditerranée*.

2. Le dépôt légal de l’internet a été instauré par la loi DADVSI de 2006. Cette mission patrimoniale est mise en œuvre par la BnF, l’INA et le CNC. La loi introduit une exception au droit d’auteur en autorisant les organismes habilités à collecter les contenus diffusés en ligne. La consultation des archives se fait uniquement sur des postes dédiés dans les espaces de recherche de la Bibliothèque ainsi qu’au sein d’un réseau de bibliothèques partenaires. La reproduction des archives pour diffusion n’est pas autorisée sauf accord des ayants droit. Les collections d’archives web de la BnF remontent à 1996 pour les plus anciennes. Fin 2024, elles représentaient 59 milliards d’URL et 2,5 petaoctets de données.

3. Depuis 2021, le BnF DataLab lance un appel à projet à destination des chercheurs souhaitant travailler, à titre individuel ou collectif, sur les collections et données numériques de la BnF et expérimenter de nouvelles méthodologies ou outils informatiques.

archivistes, documentalistes, bibliothécaires, informaticiens et historiens, mais aussi les conservateurs, ethnologues et équipes techniques impliquées dans la production de certains des sites⁴. Le 19 décembre 2024, une journée de restitution au Centre de conservation et de ressources du Mucem a permis de présenter les enjeux et les résultats de cette entreprise de remédiation. Le présent article en est une synthèse. Nous souhaitons donc remercier tous les participants et participantes qui ont contribué à cette riche discussion⁵ dont nous espérons qu'elle puisse fournir outils et méthodologies à d'autres institutions patrimoniales et culturelles confrontées aux mêmes enjeux.

La collection *Recherches ethnologiques*

Contexte institutionnel

La création du Musée des civilisations de l'Europe et de la Méditerranée (Mucem) (« Décret N° 2013-157 Du 21 Février 2013 Portant Création de l'Établissement Public Du Musée Des Civilisations de l'Europe et de La Méditerranée (MuCEM) » 2013) représente le premier véritable transfert d'un musée national en région. Le nouveau musée va mettre en œuvre un projet résolument multidisciplinaire et élargir les collections au bassin méditerranéen, ce qui sera accompli grâce auxancements de nouvelles enquêtes-collectes ethnologiques. Son histoire est cependant bien plus ancienne, car l'institution prend la suite du Musée National des Arts et Traditions Populaires (MNATP)⁶.

Devant la baisse de fréquentation de l'établissement, sa fermeture est envisagée dès les années 1990, et se concrétise en 2005, laissant place à un projet de transfert des collections dans un futur musée portant sur l'Europe et la Méditerranée. Malgré les contours incertains du projet, les enquêtes-collectes⁷ s'élargissent durant les dernières années du MNATP à ce nouveau périmètre géographique. Durant cette période de transition, le Mucem a une existence administrative sans disposer de son bâtiment d'implantation définitif. Entre la date

4. Les auteurs voudraient ici remercier tout particulièrement Jean-Pierre Dalbéra et Mohan Danabalou, de la Mission recherche et technologie du Ministère de la Culture, initiateurs de cette collection, et dont les témoignages ont été une précieuse source pour ce projet.

5. Par ordre alphabétique : Jérôme Acquafresca (Mucem, DSI), Sara Aubry (BnF, DSI), Marie-Charlotte Calafat (Mucem), Géraldine Camille (BnF), Sarah Dietz (Mucem), Simon Dumas-Primbault, professeur junior CNRS de science ouverte en SHS (OpenEdition), Alexandre Faye (BnF), Jasmine Fejzovski (stagiaire Mucem), Sophie Gebeil, maîtresse de conférences en Histoire et vice-présidente déléguée aux Sciences de l'Homme et de la Méditerranée (AMU, UMR TELEMMe), Mireille Jacotin, conservateur en chef, responsable du Pôle Vie Publique (Mucem), Édouard de Laubrie, chargé de collections et de recherches, responsable du Pôle Agriculture et alimentation (Mucem), Marie-Barbara Le Gonidec, ingénieure d'étude du ministère de la Culture, laboratoire Héritages (UMR 9022, Cergy Paris Université – CNRS – ministère de la Culture), Audrey Pety, designer (Laboratoire d'innovation numérique de la CNIL), Charles Riondet (Mucem), Chloé Sussan-Molson (stagiaire Mucem), Filippo Vancini, responsable d'exploitation multimédia (Mucem), Yannick Vernet, responsable du fablab et des projets numériques (ENSP Arles). Nous remercions également Laurence Duffès et Manuel Del Cerro (Bibliothèque de l'Alcazar, Marseille).

6. Le musée national des Arts et Traditions populaires a été créé en 1937. Il hérite des collections de la section française du Musée d'Ethnographie du Trocadéro, fondé à la fin du XIX^e siècle (Segalen 2019).

7. « L'enquête-collecte est un processus de construction de matériel d'étude du musée : des enquêtes ethnographiques sont commandées pour observer un fait de société ; des témoignages matériels et immatériels sont alors collectés sous formes d'entretiens filmés, de photographies ou d'objets ; l'ensemble est alors analysé systématiquement avant de rejoindre les fonds du musée. » Voir Ferloni et Sitzia (2022, n. 11).

officielle de sa création, en 2005, et son ouverture au public, en juin 2013, le musée se déploie en ligne en proposant de nouveaux sites web thématiques. Véritables projets éditoriaux, ils ont été durant les quatre premières années de fermeture un moyen privilégié de valorisation des collections du musée et de la recherche, notamment des enquêtes-collectes.

| Nom | Date | Contexte | Intervenants scientifiques |
|--|--------------------|---|---|
| L'olivier, trésor de la Méditerranée | 2005 | Projet européen | Edouard de Laubrie (Mucem) |
| Hip-Hop, art de rue, art de scène | Juin 2005 | Exposition du même nom au Fort St Jean | Claire Calogirou (CNRS) |
| Café, Cafés | 2005 | Partenariat avec Malongo en prévision d'une exposition | Mireille Jacotin (Mucem) |
| Les voyages du verre | 2006 | Valorisation de l'enquête collecte « Le souffle et la marque » | Denis Chevallier (Mucem) |
| Trésors du quotidien Europe et Méditerranée | printemps-été 2007 | Exposition de préfiguration du même nom, organisée au Fort Saint Jean | Denis-Michel Boëll (Mucem), Claire Calogirou, Emilie Girard, Florence Pizzorni (Mucem) |
| Cornemuses d'Europe et de Méditerranée | 17 mars 2007 | Valorisation d'une enquête-collecte mettant en avant la composante sonore | Marie-Barbara Le Gonidec (Mucem) |
| Petites Arménies d'Europe et de Méditerranée | 2007 | Réalisé dans le cadre de l'Année de l'Arménie en France « Arménie, mon amie » | Claire Mouradian (CNRS), Florence Pizzorni-Itié (Mucem) |
| Féminin-Masculin | 2008 | Valorisation des enregistrements de l'enquête-collecte « Mariages » et de plusieurs entretiens de l'enquête-collecte « SIDA » | Collectif de 14 personnes, composé d'universitaires, de conservateurs et de conservatrices. |

Les huit productions furent hébergées sur le site Ministère de la Culture, où elles étaient rassemblées au sein de la collection *Recherches ethnologiques*. Techniquement obsolètes, car ils utilisaient abondamment la technologie Flash d'Adobe, ces sites ont été mis hors ligne fin 2020.

Une expérience de préfiguration virtuelle ?

Les sites web de la collection *Recherches ethnologiques* font partie intégrante d'un projet qui, selon Yannick Vernet, a pour « objectif politique de montrer que le futur Mucem n'était pas qu'un musée de faux et de marteaux, qu'il y avait une collection d'une grande richesse et une utilisation des moyens les plus modernes et les plus performants de l'époque pour les montrer »⁸. À ce titre, le premier site web réalisé, *L'olivier, trésor de la Méditerranée* est emblématique puisqu'il porte sur un élément de culture alimentaire caractéristique des mondes méditerranéens. Il marque ainsi le début d'une entreprise de redéfinition du projet scientifique. Les sites web ne sont pas les seuls moyens mis en œuvre par le Mucem pour maintenir son activité. Durant cette période, plusieurs expositions temporaires sont proposées au Fort Saint-Jean de Marseille et certaines deviendront même itinérantes et se déclineront au travers d'une exposition virtuelle. Le musée cherche à favoriser l'interactivité en s'appuyant sur l'innovation numérique. Pour l'exposition *Hip-Hop, art de rue, art de scène*, un jeu interactif est proposé aux visiteurs, tandis que l'exposition *Trésors du quotidien* est accompagnée par un jeu télévisé et des audioguides à télécharger. Ainsi, les sites s'inscrivent dans un dispositif global de dématérialisation du musée, qui remédie à son absence d'ancrage et s'inscrit dans un projet de redéfinition de son identité.

Il est intéressant de noter que l'équipe multimédia chargée de la conception des sites est aussi celle qui gère ces installations numériques et ces productions audio et vidéo. Il existe ainsi de fait une porosité entre ces différentes activités. Le choix des sujets abordés par les sites web dépend tout à la fois des initiatives des chercheurs, de leur intérêt pour ces formats multimédias mais aussi de la programmation culturelle comme dans le cas du site *Loïn de l'Ararat* développé dans le cadre de l'année culturelle *Arménie, mon amie (2006-2007)*⁹.

Pour mettre en place ce projet éditorial, qui implique un accès gratuit pour les internautes au site, le Mucem bénéficie du soutien de la Mission de la Recherche et de la Technologie (MRT) du Ministère de la Culture. La MRT finance et héberge la collection sur son *Portail multimédia* au côté d'autres projets similaires, en particulier dans le domaine de l'archéologie. Les sites sont conçus comme des projets multimédias et présentent de fortes similarités avec les CD-ROM éducatifs et culturels des années 1990 et 2000, support utilisé par de nombreux musées (Lavigne 2005). Ces similarités se retrouvent jusque dans les technologies utilisées¹⁰, ce qui donne rétrospectivement une forte homogénéité à la collection *Recherches ethnologiques*.

L'essor du Web 2.0 à la fin de la décennie 2000 modifie néanmoins les pratiques et les attentes des internautes. Les blogs et les réseaux sociaux numériques sont alors en plein essor. L'équipe multimédia du Mucem tente d'introduire ces innovations malgré un modèle de production contraint par les critères imposés par la MRT, qui doit au final intégrer ces sites sur son portail. Néanmoins, cette influence des nouvelles pratiques numériques peut se lire en filigrane comme dans le site *Loïn de l'Ararat*, qui intègre un blog consacré à différentes personnalités

8. Citation faite à partir de la transcription de la journée d'étude du 19 décembre 2024.

9. Les années ou saisons culturelles sont des manifestations organisées par les ministères de la Culture et des Affaires étrangères. Elles visent à faire découvrir les richesses culturelles du pays invité en France et réciproquement à présenter celles de la France dans ce même pays.

10. Flash était le pendant web de Director pour les CD-ROM, soit l'outil auquel plusieurs membres de la MRT et Filippo Vancini, responsable exploitation multimédia pour le Mucem, avaient été formés.

de la communauté arménienne de Marseille. Même si l'équipe multimédia aurait souhaité une intégration plus forte des sites dans le nouvel écosystème numérique, notamment en s'ouvrant aux réseaux sociaux, les sites restent globalement fermés. Les sitographies proposées présentent d'autres sites du même univers thématique, mais elles fonctionnent plutôt comme des bibliographies web fermées sur elles-mêmes¹¹. Là où l'entrée du numérique au musée s'est articulée à un idéal de participation des visiteurs à travers une offre interactive (Vidal 2018), le projet *Recherches ethnologiques*, proche dans sa structure du CD-ROM, ne s'inscrit pas encore tout à fait dans cette logique. Finalement, les sites s'apparentent à des objets, car ils ne sont pas destinés à être mis à jour, même si certains le seront, comme *Cornemuses d'Europe et de Méditerranée*. En ce sens, ils n'ont pas de dimension temporelle, comme peuvent l'avoir un site d'actualité ou un blog, dont les pages d'accueil vont plutôt s'apparenter à un flux. Ce repli des sites sur eux-mêmes est une caractéristique essentielle, qui facilite aujourd'hui leur traitement documentaire et leur remédiation, car ils présentent une forte unité documentaire et de temps.

L'expérience originale que représente la mise en ligne des huit sites de la collection *Recherches ethnologiques* prendra fin au moment de la création de l'Association de préfiguration du Mucem et du lancement du chantier de construction. Alors qu'ils occupaient une place centrale dans la stratégie communicationnelle du musée, ils vont s'effacer sans engendrer de réflexion sur leur place à moyen terme. Ils auront été une expérience temporaire, un bref épisode, précédant le transfert du musée à Marseille.

Traitement archivistique du corpus

Dans un premier temps, l'intérêt pour les archives web de la collection *Recherches ethnologiques* était guidé par une volonté de récupération de données scientifiques présentes sur ces sites et susceptibles d'enrichir les bases documentaires du Centre de conservation de ressources du Mucem. Ce besoin concernait plus particulièrement le site *Cornemuses*, dont le contenu scientifique était toujours d'actualité et qui s'apparentait à une véritable base de données en ligne décrivant ces instruments conservés par le musée. Marie-Barbara Le Gonidec, chercheuse impliquée dans la direction scientifique de ce site, avait alerté le Mucem sur le fait que les données en ligne étaient plus riches et méritaient d'être récupérées.

L'idée d'une remédiation des sites à partir de l'archive web émergea après une séance de consultation à la BnF en présence de Marie-Barbara Le Gonidec, ce qui conduisit à la conception du projet « Virtuel Mucem ». Plus qu'une copie de l'archive ou une extraction de données, le projet s'est donné pour objectif de mettre en œuvre une véritable intégration des archives web dans les bases documentaires. Cette intégration devait permettre la consultation et la navigation selon un principe de rejouabilité, la recherche de contenus grâce à une indexation et embarquer une contextualisation des sites. Il fut rapidement évident que la remédiation de l'archive web exigerait un travail documentaire complet pour pouvoir décrire les sites avec précision, les contextualiser et remédier à d'éventuelles lacunes intrinsèques, notamment l'absence de certaines vidéos ou l'impossibilité d'afficher certaines animations 3D.

11. De manière générale, les sites ont très peu de liens hypertextes et il n'y en a pas sur les sitographies, ce qui ne permet pas une navigation fluide vers ces ressources signalées.

Or, une masse importante de documents avait été produite à l'époque de la réalisation des sites. Ces fichiers étaient conservés sur les serveurs numériques du Mucem et n'avaient fait l'objet d'aucun traitement archivistique, si bien qu'il était difficile de savoir ce qu'ils recouvraient. Cet ensemble fut désigné sous le terme de « vrac numérique ». Néanmoins, dès les premiers sondages, cette documentation s'avéra très riche. Elle contenait des fichiers techniques (des sauvegardes des sites, des fichiers de programmation pour créer des gabarits, etc.), des fichiers liés à la gestion de projet (échanges de courriels, comptes-rendus de réunion, communications autour des sites), ainsi que les fichiers liés aux recherches documentaires.

Le programme de travail combina plusieurs approches : un traitement archivistique du vrac numérique, un enrichissement des bases documentaires locales et enfin la constitution d'archives orales grâce à la réalisation d'une série d'entretiens auprès des acteurs de l'époque. Ces tâches furent menées par les équipes du Centre de conservation et de ressources, qui accueillit deux stagiaires en renfort sur cette mission.

Les témoins contactés pour les entretiens ont été choisis sur des critères essentiellement pratiques : il s'agit de personnes ressources identifiées ayant participé à la production des sites web et pouvant apporter un éclairage sur le contexte et des conseils pour retrouver la documentation produite à l'époque. Différents types d'acteurs ont été impliqués : agents du service multimédia, au niveau décisionnaire ou opérationnel, et auteurs des contenus. Les entretiens ont permis de comprendre le contexte de production propre à chaque site. Pour autant, chacun d'entre eux se rattache à une histoire plus générale, celle de l'institution, de la muséographie, de la médiation culturelle, de l'évolution des techniques, d'acculturation au numérique et de développement de l'internet grand public au tournant des années 2000. Le projet n'a pas cherché à développer plus en avant ces pistes de recherche, qui aurait nécessité de multiplier les entretiens auprès d'autres acteurs (personnel de recherche, prestataires de service, partenaires du musée, etc.) et d'établir des comparaisons avec d'autres collections soutenues par la Mission recherche et technologie du Ministère de la Culture.

Pour commencer son travail, l'équipe consultait les archives d'Internet Archive¹², dont les collections sont librement accessibles en ligne. En parallèle, la BnF prépara deux modes d'extraction des archives web pour les livrer à l'équipe du Mucem. Le premier envoi consistait en une extraction des fichiers présents dans l'archive (les pages HTML, les fichiers multimédia, ainsi que les feuilles de style). L'objectif était de prendre connaissance avec plus de précision du contenu de l'archive web et de le comparer avec les éléments trouvés dans le vrac numérique. Les sources multimédias pouvaient également présenter des différences liées au travail de post-production réalisé avant la mise en ligne. La seconde livraison proposait une extraction de l'archive web au format WARC (Web ARChive) pour laquelle la BnF avait réalisé un développement permettant de parser l'archive (cf. *infra*). L'utilisation d'un format standard présente un avantage essentiel, puisqu'il permet l'utilisation d'outils de consultation et de recherche développés par la communauté internationale des archivistes du web. Ce sont, en particulier, les fonctionnalités de recherche qui permirent de fouiller directement dans les sites.

12. La fondation Internet Archive, pionnière dans l'archivage du web et à l'origine du développement de la Wayback Machine (outil de consultation, qui permet une navigation entre les pages archivées par lien hypertexte, mais aussi par date de capture) propose une vaste archive de sites internet librement accessible en ligne.

La présence des numéros d'inventaire sur le site *Cornemuses* facilita le travail d'intégration des données dans les bases documentaires.

À l'issue de cette phase, les archives numériques classées reçurent une cote et intégrèrent le système de gestion des archives de l'établissement. Ces instruments de recherche constituent à la fois l'aboutissement de ce travail et une forme de reconnaissance de l'importance de la collection *Recherches ethnologiques* dans l'histoire du musée.

La question de la vérification des droits d'auteurs et droits à l'image initiaux, qui aurait été nécessaire dans le cadre d'une éventuelle remise en ligne sur Internet, n'a pas été approfondie, les contraintes techniques de la remédiation imposant un dispositif hors ligne dans les locaux de l'établissement (cf. *infra*). L'accès à ces archives, comme aux fichiers obtenus en complément par les archives du web, se fait donc selon les modalités d'accès aux archives publiques, le Mucem en étant producteur.

Le dispositif technique de remédiation de l'archive

Produire une nouvelle archive web à des fins de remédiation à partir des collections de la BnF

Les sites de la collection *Recherches ethnologiques* ont été archivés à plusieurs reprises par les robots de la BnF entre 2007 et 2020. Cet archivage régulier a été réalisé dans le cadre des collectes courantes de la bibliothèque et génère différentes captures des sites, qui se recoupent largement puisque ces sites n'ont que superficiellement évolué entre ces dates. Concrètement, à chaque collecte, les robots de la BnF copient les contenus du site mais aussi les fichiers associés à la mise en page, des scripts ou d'autres éléments liés à l'environnement technique du site. Ces sites utilisaient largement la technologie Flash d'Adobe, aussi bien pour l'animation des sites que pour le format des vidéos mises en ligne.

Ces archives sont stockées sur les serveurs de la BnF¹³ dans des fichiers conteneurs au format WARC, objet d'une norme internationale s'appuyant sur les travaux du consortium international pour la préservation de l'internet (IIPC) (« ISO 28500:2017 Information and Documentation – WARC File Format » 2017). Pour bien comprendre la nature de l'archive web, il faut préciser ici que ce format n'est pas organisé intellectuellement. Il ne procède pas d'un travail documentaire qu'aurait mené le déposant du site, ici le Mucem en tant que producteur, ou les archivistes de la BnF. Lors des collectes courantes, plusieurs sites sont moissonnés en parallèle, jusqu'à plusieurs milliers, générant un flux de données conséquent. Le format WARC est conçu pour absorber efficacement ce flux et permettre la sauvegarde des métadonnées et des enregistrements au sein d'un même bloc d'information¹⁴. Autrement dit, au sein d'un même fichier WARC, il est possible de retrouver des blocs d'information

13. Le stockage s'effectue sur des espaces sécurisés non accessibles depuis l'extérieur afin de protéger ces contenus, qui restent soumis au droit d'auteur. Le système de préservation et d'archivage réparti de la BnF (SPAR) s'appuie sur les standards METS et PREMIS.

14. Les fichiers WARC sont structurés par bloc. Chaque bloc contient un enregistrement qui peut être une page HTML, une image, une feuille de style ou un autre type de fichier web. Les enregistrements sont précédés de métadonnées qui décrivent le fichier (comme l'URL, la date exacte de capture, le type MIME du fichier renvoyé par le serveur, sa taille, etc.).

issus de différents sites. Il est aussi possible que la capture d'un site ne soit pas complète, en fonction des liens suivis ou non par le robot.

Pour autant, il a été décidé dès l'origine du projet de conserver le format WARC pour concevoir la remédiation. Cela impliquait de produire de nouveaux fichiers WARC, dérivés des fichiers originaux, en les délimitant d'un point de vue documentaire aux sites du Mucem. Une autre option aurait été d'extraire uniquement les fichiers contenant le texte ou correspondant à des médias (image, son, vidéo, dont les fichiers Flash) et de les fournir aux équipes du Mucem, charge à elle de « reconstruire » un nouveau site ou d'incorporer ces éléments dans une base de données.

Cette proposition pouvait satisfaire certains besoins du musée, notamment la récupération d'information scientifique à partir de numéro d'inventaire d'artefacts pour enrichir les bases documentaires du musée. Cependant, elle ignorait de fait la question de la remédiation et les problématiques qui étaient justement au cœur du projet : comment assurer la remédiation de contenus critiques pour la sécurité informatique comme les fichiers Flash ? Quels regards pouvons-nous porter sur ces formes de médiations passées ? Comment transmettre cette histoire et quel dispositif narratif pourrait être mis en place ?

```

WARC/1.0
Content-Length: 54140
Content-Type: application/http; msgtype=response
WARC-Date: 2021-07-01T15:48:04Z
WARC-IP-Address: 143.126.211.212
WARC-Payload-Digest: sha1:RXMYTELTSFP5LLZZ5KT6U26I33IE5TKB
WARC-Record-ID: <urn:uuid:d602f53b-ebad-4982-b247-fbe876537df7>
WARC-Target-URI: http://www.ethnologie.culture.fr/olivier/flash/savoir\_bois.swf
WARC-Type: response

HTTP/1.1 200 OK
Date: Thu, 01 Jul 2021 15:48:04 GMT
Server: Apache-AdvancedExtranetServer
Last-Modified: Tue, 15 May 2007 13:55:41 GMT
ETag: "9887bb-d264-96d0f140"
Accept-Ranges: bytes
Content-Length: 53860
Connection: close
Content-Type: application/x-shockwave-flash

CWSBELDELxF9NULNULx9CxECxBBELTDC4KxF7/ xDA3xC3FFixC8EMAGrxCCEMIx92%Jx92 (qxB
BSHx92, x88xA8#DLEETXxA8xA0xF2zDLExF0x9CxF3x9DxFBxFE7xBDxB5xEE [ xEB6x8BxE6wxi
xE0xACEOTcBELNUL x95DLE<DC1xC2xC1DC3 [*1x80NUL.SOSO`DLENULzEM

```

FIGURE 1 – Extrait d'une animation Flash du site *L'olivier, trésor de la Méditerranée* stockée en WARC. La première et la seconde partie du bloc contiennent des informations techniques sur le WARC et la collecte des métadonnées associés à ce fichier Flash. La troisième partie correspond à l'enregistrement et embarque directement le fichier Flash, qui nécessite un lecteur adapté pour pouvoir être lu.

Conception de l’outil d’extraction et création d’un environnement de consultation sécurisé

Parallèlement au traitement archivistique et aux enquêtes orales menées par l’équipe du Mucem, les équipes techniques de la BnF ont développé un outil capable de générer les fichiers WARC dérivés. Nommé WARCD, il permet aux chargés de collections numériques de la BnF de configurer l’extraction des fichiers en sélectionnant un site, en requêtant les archives de la BnF sur une période de temps choisie et en décidant de conserver l’ensemble des captures ou seulement la dernière en date. Ce dernier critère est d’ailleurs bien adapté aux sites n’ayant que marginalement évolué dans le temps comme ceux de la collection *Recherches ethnologiques*, plutôt considérés comme des objets autonomes. Ainsi, WARCD permet de composer l’archive dérivée la plus complète possible à partir de tout ce qui a pu être collecté par la BnF. Il faut également noter que l’outil gère en entrée le format WARC, ainsi que le format ARC (ARChive) qui l’a précédé et qui a été utilisé jusqu’au début des années 2010¹⁵.

Après l’extraction pour chaque site des nouveaux fichiers WARC dérivés, une instruction fut ouverte pour définir le meilleur environnement technique de consultation. Les fichiers WARC ne peuvent pas être directement ouverts avec un navigateur. Une application est nécessaire pour en indexer le contenu et permettre de naviguer dans l’archive. Deux pistes ont été explorées :

- Le service en ligne Conifer¹⁶, qui permet de créer ses propres archives web, de les consulter et de les télécharger au format WARC. Cette solution ne demande pas d’installation locale, ce qui représente un avantage du point de vue de la sécurité informatique (pas besoin d’installer des applications qui ne sont pas maintenues et pourraient présenter des failles de sécurité). Cependant la restitution de l’archive s’est avérée incomplète, car certaines pages du WARC n’étaient pas traitées. De plus, la mise en ligne sur internet, donc le dépôt des WARC sur des serveurs extérieurs, auraient soulevé des questions juridiques sur les droits de diffusion et de réutilisation.
- L’application *open source* SolrWayback¹⁷, qui se présente à la fois comme un outil de consultation et une boîte à outils offrant de nombreuses fonctionnalités de recherche (indexation plein texte, recherche par images, nuage de mots, etc.). Installée en local au Centre de conservation et de ressources du Mucem, elle offre un affichage de qualité et permet de lancer directement des recherches au sein des archives remplissant ainsi les conditions attendues.

15. Techniquement, l’outil permet aussi de tracer l’origine des fichiers WARC dérivés en ajoutant des métadonnées de contexte lors de la génération. Il fournit un rapport permettant d’avoir la liste complète des URL des fichiers extraits et leur correspondance entre leur localisation à la BnF et leur localisation dans les WARC dérivés.

16. Conifer est une plateforme d’archivage en ligne qui permet de créer ses propres archives web à partir du web vivant, de les consulter et de les télécharger au format WARC. Elle propose aussi d’héberger des archives (dans la limite de quelques gigaoctets) et de les rendre accessibles publiquement. Elle a été créée en 2015 par l’association américaine Rhizome qui s’intéresse à la scénarisation et à la préservation de l’art numérique pour les musées. Conifer intègre un navigateur virtuel qui permet de rejouer les archives.

17. Développé par la bibliothèque royale du Danemark, SolrWayback est utilisé par de nombreuses institutions au niveau international. La BnF contribue au développement et utilise ce logiciel pour faciliter la prise en main de corpus réalisés à partir de collecte collaborative menée avec des équipes de recherche.

Enfin, les équipes informatiques de la BnF et du Mucem ont mené un travail conjoint pour mettre en place un environnement logiciel compatible avec les fichiers son, vidéo et les animations 3D contenus dans les sites. Pour rejouer ces formats multimédias, qui ne sont plus pris en charge par les navigateurs récents, il a fallu installer des versions logicielles adaptées (la version 78 de Firefox, le plugin Flash 32.0.0.71¹⁸, le Virtools plugin¹⁹ pour la 3D et VLC media player pour le son).

L'installation a pu être effectuée avec succès en vue de la journée de restitution du projet. L'expérience montre l'intérêt de l'adoption de standards et d'outils partagés pour des projets de préservation-remédiation.

18. Flash est difficilement trouvable sur le web vivant et a pu être récupéré dans les archives d'*Internet Archive*.

19. Ce plugin date de 2010.

Éditorialisation

Les limites présentées par la remédiation des fichiers WARC et des archives numériques exigeaient, du point de vue du Mucem, la mise en place d'un dispositif d'accompagnement. En effet, le poste de travail qui accueille la SolrWayback était destiné à être installé dans la salle de lecture du Centre de conservation et de ressources, salle ouverte aux chercheurs, aux agents du musée mais aussi au grand public. L'accès aux archives du web présente des spécificités auxquelles ces différents publics sont généralement peu sensibilisés : en premier lieu son caractère lacunaire et reconstruit. Il s'agissait donc de proposer des éléments de compréhension du contexte historique et institutionnel, une présentation succincte de la démarche de restitution, mais aussi en détaillant pour chacun des sites ce que l'archive donne ou non à voir. Concrètement, l'éditorialisation se présente sous la forme d'un site statique²⁰, déployé en interne, depuis lequel il est possible d'accéder aux archives web pour retrouver le plaisir de parcourir ces objets du passé.



FIGURE 2 – Page d'accueil du site de remédiation *Virtuel Mucem*

Pour chacun des sites, une page de synthèse présente un bref rappel du sujet, parfois le caractère obsolète de certains contenus, puis dresse un état des lieux sommaire de l'état de l'archive virtuelle : quelles en sont les lacunes (textes ou médias manquants), les requêtes possibles pour accéder à certains contenus, les spécificités par rapport aux autres sites de la collections ? Et quel est le degré de rejouabilité des animations Flash ? S'y ajoute un

20. L'équipe a développé la maquette du site avec la version en ligne gratuite de l'application Notion. Louis Moreau-Gaudy, ingénieur Huma-Num au BnF DataLab, a apporté ses conseils et son appui technique pour la génération automatique du site statique réalisé avec Eleventy (<https://www.11ty.dev>). L'hébergement du projet sur la plateforme GitLab de Huma-Num facilite l'évolution des contenus et leur mise à jour.

lien vers les ressources et médias complémentaires issus des archives numériques du Mucem. L'éditorialisation est pensée pour permettre aux lecteurs de considérer ces archives numériques pour ce qu'elles sont : des archives web témoignant d'un dispositif numérique passé, des traces laissées par l'histoire d'une institution et de ses pratiques de médiations, ainsi que l'état de la recherche ethnologique à un moment donné.

Présentation de deux sites de la collection

Cornemuses d'Europe et de Méditerranée



FIGURE 3 – Capture de la page d'accueil du site *Cornemuses* le 24 octobre 2007 dans les archives de l'Internet de la BnF (<http://archivesinternet.bnf.fr/20071024201329/http://www.cornemuses.culture.fr/>)

Le site *Cornemuses d'Europe et de Méditerranée* a été réalisé en collaboration avec l'ethnomusicologue Marie-Barbara le Gonidec. Il se compose d'une partie thématique présentant la place de l'instrument dans les sociétés sous l'angle historique et culturel, et d'une base de données correspondant à un inventaire des cornemuses conservées. La collection avait été enrichie par une enquête-collecte menée dans le bassin méditerranéen.

Marie-Barbara le Gonidec s'était fortement investie dans la réalisation du site et continua à alimenter ponctuellement la partie base de données dans les années qui suivirent sa mise en ligne. Le site inclut des vidéos et des enregistrements sonores, essentiels pour transmettre les connaissances, là où l'exposition physique offre souvent moins d'opportunités.

Féminin-Masculin, histoires de couples et construction du genre



FIGURE 4 – Capture de la page d'accueil du site *Féminin-Masculin* du 28 avril 2010 dans les archives de l'Internet de la BnF (<http://archivesinternet.bnf.fr/20100418042601/http://www.femininmasculin.culture.fr/>)

Publié en 2008, « Féminin-Masculin, histoires de couples et construction du genre » est le dernier site de la collection *Recherches ethnologiques*. L'idée du site vient de la volonté de valoriser un ensemble de 22 DVD issus de l'enquête-collecte *Mariage*, le web est vu comme la meilleure manière de présenter ces vidéos et le travail de recherche. Le site propose de découvrir différents couples et met en avant les rites de passage, religieux ou civiques, sous l'angle de leur fonction d'ordonnement de l'identité genrée et des rapports sociaux entre les sexes. Le site intègre également des matériaux issus de l'enquête-collecte *SIDA*.

Ainsi, à l'instar du site *Cornemuses d'Europe et de Méditerranée*, *Féminin-Masculin* permet la mise en valeur de travaux ethnologiques pour lesquels une logique d'« objet-témoin » ne suffirait pas : les vidéos et les témoignages des couples, fruits d'entretiens semi-directifs, sont au cœur du projet. Le site s'apparente à une exposition virtuelle. Il montre aussi la volonté du nouveau musée d'aborder des thématiques sociétales en abordant aussi bien les traditions folkloriques que les pratiques contemporaines. Il servira d'inspiration pour l'exposition au *Bazar du genre* de 2013, proposée lors de l'inauguration du Mucem.

Une journée de restitution pour redécouvrir les sites dans leur contexte

Une des réussites du projet Virtuel Mucem est la création d'une réelle « adhésion au projet »²¹, quels que soient les profils professionnels, l'institution de rattachement et le niveau d'implication des acteurs participants. La journée de restitution les a réunis le 19 décembre 2024, au Centre de Conservation et de Ressources du Mucem, en présence de celles et ceux qui contribuèrent à la réalisation des sites de la collection *Recherches ethnologiques* : Mireille Jacotin, Édouard de Laubrie, Marie-Barbara Le Gonidec, Filippo Vancini et Yannick Vernet²².

La journée proposait de découvrir le dispositif de remédiation à travers une démonstration commentée et un retour sur les problématiques documentaires et techniques auxquelles se sont confrontées l'équipe projet. Deux tables rondes ont rythmé les échanges : la première revenant sur le contexte de production des sites, et la seconde questionnant l'usage et la reproductibilité de l'expérience en présence de l'historienne Sophie Gebeil et de la designeuse Audrey Pety. Cette journée de restitution a d'abord été un plaisir, celui de retrouver des sites aujourd'hui disparus dans les espaces du Mucem et de pouvoir les parcourir à nouveau. Cette « ancienne version »²³ (Jérôme Acquafresca), qui témoigne de l'histoire du Mucem, a été une découverte pour certains participants. Pour d'autres, c'était le plaisir de « réveiller un peu »²⁴ (Marie-Barbara Le Gonidec) des sites élaborés en lien avec les collectes de terrain.

À partir de la remédiation, c'est effectivement une histoire institutionnelle qui apparaît dans sa complexité. En effet, les sites de la collection *Recherches ethnologiques* ont porté une « revendication politique forte »²⁵ (Edouard de Laubrie). Ils ont donné à voir le résultat des premières enquêtes de terrain faites hors de France, ainsi que l'entrée d'objets contemporains dans les collections, comme le skateboard. Leur mode de production sériel et l'utilisation de la technologie Flash permettent d'établir une filiation avec les CD-ROM culturels des années 1990. « Les sites sortaient à une moyenne [d'un] chaque 3-4 mois, produits l'un après l'autre » (Filippo Vancini). Pourtant, il y avait là aussi la revendication d'un droit à tester des « choses

21. Toutes les citations sont faites à partir de la transcription de la journée d'étude. Les propos repris ici sont ceux de Jérôme Acquafresca : « le pouvoir d'adhésion, quand ce projet est arrivé, ça a été un peu comme une bouffée d'oxygène dans notre quotidien ». Au moment de faire référence à cette journée d'étude, le nom de l'interlocuteur sera ajouté directement entre parenthèses dans le texte.

22. Cette journée faisait suite à la journée d'étude « La (re-)médiation des collections de la BnF et du Mucem. Les multiples matérialités des fonds documentaires et les pratiques qu'elles révèlent » qui s'est tenue le 16 décembre au MucemLab avec la participation de Simon Dumas-Primbault.

23. « Moi qui ai connu le Mucem dans sa nouvelle version, pas dans l'ancienne, j'apprécie de voir ce contenu qui fonctionne plutôt bien, il faut le reconnaître. Mais c'est vrai qu'on n'imagine pas tout ce qui se cache derrière pour pouvoir le faire fonctionner. ».

24. « Je suis vraiment contente si on peut réveiller un peu ce site. Pour l'ethnomusicologie, un instrument de musique, ce n'est pas un objet comme un autre. Il faut qu'on l'entende, si possible qu'on le voit être joué. Ça a été une chance extraordinaire de pouvoir montrer, mettre des vidéos. ».

25. « Au tout début des années 2000, avec l'arrivée de Jean-Pierre Dalbéra, on découvre les nouvelles potentialités de l'informatique, c'est une vraie révolution mentale. Il y avait aussi une revendication politique forte, alors que le projet du Mucem avait du mal à émerger avec une opposition forte au sein de l'institution. Ces sites internet étaient l'occasion de montrer, de donner à voir les premières enquêtes de terrain faites en dehors de la France. »

expérimentales »²⁶ avec l'arrivée du web 2.0 (Yannick Vernet). Ce tournant numérique inspiré par l'émergence du web social rejoue le tournant social et communicationnel²⁷ des musées en le projetant sur ces nouveaux espaces virtuels. « L'approche scientifique est extrêmement importante et intéressante, mais il n'y aura pas de rencontre avec le public si on ne s'intéresse pas aux usagers » (Yannick Vernet). L'expérience de la remédiation convoque le souvenir des collectes de terrain en particulier avec le site *Loin de l'Ararat* qui garde une place particulière auprès de « toute cette communauté arménienne de Marseille, qui, aujourd'hui, nous disent : "on aimerait bien revoir ces images de l'aventure" » (Yannick Vernet).

Au-delà des témoignages des acteurs, la question de la réception de la collection des années 2000 reste mal connue, contrairement au premier site internet du MNATP, qui avait fait l'objet d'une étude de réception en 1997 (Châtillon et Cachin 1997). Les statistiques de fréquentation n'ont, par exemple, pas été identifiées dans le vrac numérique. Une approche plus générale, à l'échelle de l'ensemble des sites soutenus par la MRT, serait nécessaire pour approfondir ces points.

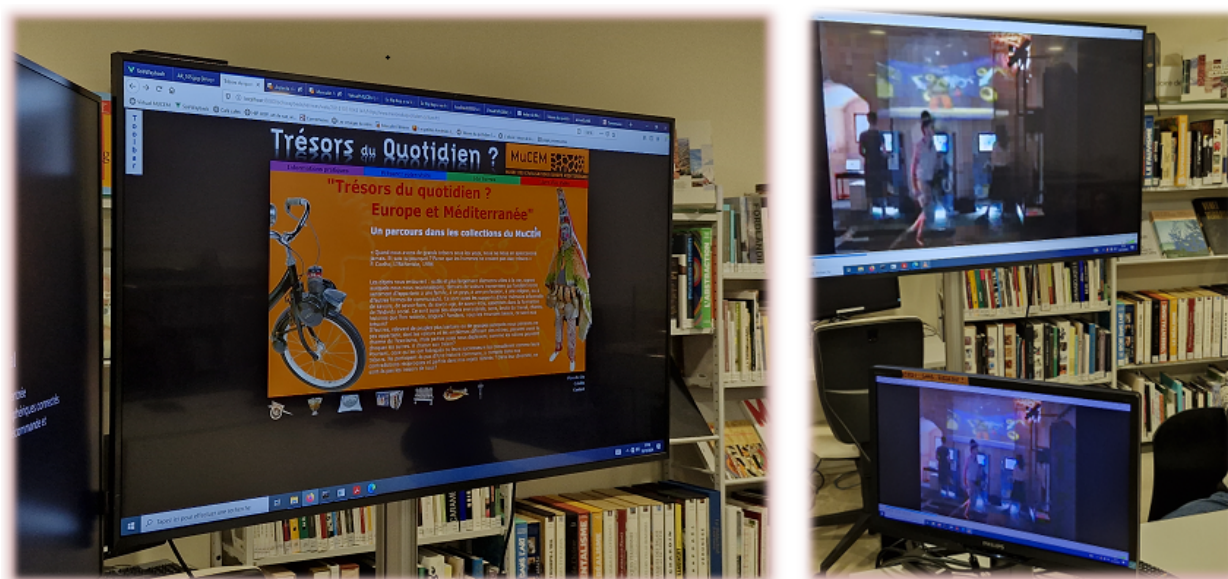


FIGURE 5 – Remédiation des sites *Trésors du quotidien* et *Hip-Hop, art de rue, art de scène*, lors de la journée de restitution au Centre de Conservation et de Ressources du Mucem.

26. « On était vraiment dans le Web 2.0, c'était les années 2000, il y a plein de choses qui se passaient. On voulait insuffler tout ça aussi dans les sites internet. Mais on ne pouvait pas faire tout ce dont on avait envie, car les sites dépendaient du Ministère de la Culture. On a pu innover timidement. Pour Arménie, on a fait une espèce de faux blog avec toute la production multimédia de Filippo. On a fait beaucoup de choses expérimentales dans le cadre de *Trésors du quotidien*, et si besoin on déportait les contenus comme pour le projet avec TV5 monde. »

27. Le terme « communicationnel » est ici employé dans le sens que pouvait lui donner Jean Davallon (1992, 117), à savoir comme une composante de la mission sociale du musée. « Ce processus indique une approche du public par l'institution muséale qui n'est ni seulement communicationnelle (centrée sur la transmission du savoir), ni seulement économique (en termes de clientèles), mais va plutôt dans le sens de sa fonction au sein de l'espace public, pour un public qui va, en ce cas, visiter l'exposition comme on va voir un spectacle ».

La discussion est revenue sur le travail de documentation de l'archive, qui a abouti à la contextualisation et à l'éditorialisation des archives des sites. Le projet a montré que cette phase est essentielle. Il ne s'agit pas ici de remettre en ligne les contenus. Les intervenants soulignent d'ailleurs que ces sites sont datés. Par exemple, le site *Café, Cafés* constituait à la fois une sorte de « *teaser* » pour annoncer l'exposition, mais faisait aussi le point sur « l'état de la connaissance liée au café » (Mireille Jacotin). Or les connaissances rapportées ne sont plus à jour, car « c'était avant l'interdiction européenne de fumer dans les espaces publics, en particulier dans les cafés ; avant la diffusion du café sous forme de dosette ; avant la généralisation des chaînes commerciales ; avant la vogue, aussi, des cafés chicha » (Mireille Jacotin).

L'éditorialisation vient aussi compléter les « trous »²⁸ de l'archive (Alexandre Faye), qui peuvent générer une certaine déception dans un dispositif de remédiation basé sur la rejouabilité et qui donne l'illusion d'être plongé dans le web de l'époque. Les pages éditoriales viennent ainsi signaler les lacunes comme les vidéos manquantes. Elles constituent un avertissement sur la complétude de l'archive web et peuvent faire le lien avec d'autres archives et collections, qui sont issues des bases documentaires du Mucem et de la conservation des matériaux de terrain. Les pages éditoriales vont ainsi introduire l'archive web et donner un sens à la remédiation en expliquant le contexte de production des sites et en se proposant également comme notice d'utilisation.

La reproductibilité de cette forme de remédiation des archives du web n'est pas univoque. Dans le cas présent, les sites ont tous une unité très forte, qui permet de les traiter comme des objets. On peut retrouver des similitudes avec les sites commémoratifs, les sites créés spécifiquement pour un événement, comme les Jeux olympiques ou une campagne électorale. Cela permet, dans une certaine mesure, de faire abstraction de la dimension temporelle. Cependant, dans d'autres cas, le projet peut porter sur des contenus issus du web social où d'autres logiques vont prévaloir : celle du réseau, du flux et de l'algorithme. Ce sont justement ces cas de figure qui intéressent désormais les chercheurs travaillant sur des terrains numériques. Pour les musées, cette expérience de remédiation pose la question de la capacité de l'institution à préserver les nouvelles innovations, comme la réalité augmentée. Ces dispositifs passent par des applications qui n'entrent pas dans les collectes du web menées par la BnF. Néanmoins, il existe d'autres moyens d'en conserver des traces. Certains intervenants auraient d'ailleurs souhaité pouvoir faire des captations des sites de la collection *Recherches ethnologiques* avant leur décommission. Cette problématique est fortement ressentie par l'ensemble des participants, car l'innovation est au cœur de la relation entre le musée et son public. « Alors que les discours de la promesse associés aux technologies numériques valorisent généralement une innovation de rupture, il faut dire, au contraire, à quel point celles-ci sont banalisées dans les attentes des visiteurs vis-à-vis de l'institution muséale – parce qu'elles le sont plus généralement dans l'ensemble de l'offre culturelle » (Bordeaux, Poulot, et Triquet 2020, paragr. 3). Il est difficile, aujourd'hui, de savoir quelles traces laisseront ces nouvelles médiations si déterminantes dans l'histoire de la muséologie.

28. « Oui il y a des trous et ça fait partie du jeu. J'ai envie de dire que ça fait partie de la déception qu'il y a besoin d'éprouver face aux archives du web pour mieux comprendre ce dont il s'agit. On joue sur la rejouabilité, donc on joue un peu sur une sorte d'illusion. Il y a le côté magique de retrouver des choses qui ne sont plus en ligne, qui sont là, qu'on peut remettre un petit peu dans leur contexte. »

Les interventions de Sophie Gebeil et Audrey Pety permettent de présenter et d’imaginer d’autres formes de remédiation. Dans les deux cas, les chercheuses ont été confrontées à la question de la curation d’archives web autour d’un sujet d’étude. Dans le cadre de sa thèse sur *La fabrique numérique des mémoires de l’immigration maghrébine sur le web français (1999-2014)*, Sophie Gebeil a été chercheuse associée à la BnF. C’est dans ce cadre qu’elle a publié un parcours guidé des archives du web sur ce sujet. Il s’agit d’une sélection organisée de captures ou de sources utilisées dans le cadre de son travail et regroupées par thème pour en donner une présentation pédagogique. Issus d’un travail scientifique, les parcours guidés sont des moyens d’entrée dans les collections d’archives du web de la BnF et constituent ainsi une forme de médiation originale. Le travail d’Audrey Pety réalisé à l’occasion de son mémoire d’étude a porté sur une remédiation d’archives web relatives à l’adoption du mariage pour tous en 2013. La designeuse a imaginé un outil de curation permettant d’organiser les captures comme des artefacts, qui seraient ensuite utilisées dans le cadre d’une exposition. L’approche ne met pas en avant la rejouabilité de l’archive mais plutôt l’expérience du public au contact de ses images nativement numériques qui ont largement circulé et qu’il convient de pouvoir montrer et contextualiser.

Pistes pour une reproductibilité

Le projet du Mucem pourrait servir d’exemple à d’autres institutions, mais sa réitération ne serait pas forcément évidente.

L’organisation des sites de la collection *Recherches ethnologiques*, qui restent globalement antérieurs au web 2.0 dans leur conception, malgré des innovations reprises par leurs concepteurs à l’époque, a facilité leur extraction et leur remédiation. La présence de fichiers Flash a obligé les équipes à prévoir un dispositif local spécifique pour des raisons de sécurité informatique, mais n’a pas freiné le projet. Or, les sites récents peuvent être plus complexes à préserver, car ils embarquent des technologies et des contenus hébergés en dehors du site lui-même. C’est fréquemment le cas pour les supports multimédia comme la vidéo, qui peut provenir d’une chaîne Youtube gérée par le producteur du site. Les solutions techniques d’identification et de *repackaging* de ces contenus ne seraient donc pas similaires. La remédiation des sites pose aussi des questions juridiques, car les archives du web restent soumises au droit d’auteur. La remédiation d’un site implique d’obtenir un accord des ayants droit, ce qui n’était pas nécessaire dans le cas du projet Virtuel Mucem.

Les archives du web sont par essence une forme de remédiation essayant de donner l’illusion d’une représentation exacte des sites passés. Cette simplicité de principe cache en fait une complexité de ce matériel, qui peut s’avérer lacunaire ou au contraire redondant et qui nécessite de maintenir des infrastructures obsolètes pour pouvoir continuer à lire certains formats. Ces difficultés peuvent freiner les chercheurs, qui abordent plutôt la source à travers la notion de corpus. Il y a des points communs entre la production d’une remédiation comme dans le cas du projet Virtuel Mucem et la constitution de corpus à partir d’un terrain numérique passé ou présent. La méthodologie de constitution reste centrale, tout comme la pédagogie, si l’on souhaite transmettre ou exploiter la source. Le projet Virtuel Mucem montre que la remédiation présente toujours des biais inhérents au choix de traitement. Il montre aussi que la complémentarité entre le web vivant et le web archivé observée dans de

nombreux projets de recherche se retrouve ici lorsque des sources complémentaires issues des bases documentaires du Mucem viennent compléter l'archive.

Conclusion

Les institutions culturelles, comme le Mucem, se sont emparées des nouvelles technologies très rapidement pour proposer des dispositifs innovants au public sous des formes très variées (Magro et Pesquer 2016). Une fois fermés, les contenus de ces sites ou bases de données sont devenus inaccessibles, alors même que ces formes innovantes ont connu parfois un taux de fréquentation important. Les institutions culturelles se posent alors la question de leur préservation et de leur valorisation grâce à la remédiation. Ces formes de remédiation peuvent être multiples et restent à inventer. Les parcours guidés proposés par la BnF en sont une illustration, tout comme les propositions d'Audrey Pety qui valorisent l'interaction avec le public, entre émotion et contextualisation.

Les résultats de Virtuel Mucem ont aujourd'hui intégré le corpus des ressources documentaires du musée. Présentés sur un poste dédié en salle de lecture, ils sont librement accessibles au grand public comme aux chercheurs qui le souhaiteraient. Mais au-delà de l'échelle d'un établissement, la démarche se veut surtout une expérimentation, technique et documentaire, qui ouvre des pistes pour d'autres projets de ce type. Il s'agira bien entendu, pour ceux qui s'engageraient dans un chemin similaire, de décliner ses principes en fonction des spécificités locales. Nous espérons cependant avoir posé quelques jalons utiles.

Bibliographie

- Bordeaux, Marie-Christine, Dominique Poulot, et Éric Triquet. 2020. « Éditorial ». *Culture & Musées. Muséologie et recherches sur la culture*, n° 35 (juin). Université d'Avignon et des Pays de Vaucluse:7-10. <https://doi.org/10.4000/culturemusees.4373>.
- Châtillon, Élisabeth, et Françoise Cachin. 1997. *Bilan et observation d'un site sur Internet. Traditions de Noël en France et au Canada*. Paris : Direction des musées de France.
- Davallon, Jean. 1992. « Le musée est-il vraiment un média ? » *Publics et Musées* 2 (1):99-123. <https://doi.org/10.3406/pumus.1992.1017>.
- « Décret N° 2013-157 Du 21 Février 2013 Portant Création de l'Établissement Public Du Musée Des Civilisations de l'Europe et de La Méditerranée (MuCEM) ». 2013.
- Ferloni, Julia, et Emilie Sitzia. 2022. « Quand le musée de société donne du pouvoir. Enjeux de la construction participative d'expositions au Mucem ». *Culture & Musées. Muséologie et recherches sur la culture*, n° 39 (mai). Université d'Avignon et des Pays de Vaucluse:135-62. <https://doi.org/10.4000/culturemusees.7888>.
- Gebeil, Sophie, et Jean-Christophe Peyssard, éd. 2024. *Exploring the Archived Web During a Highly Transformative Age. Proceedings of the 5th International RESAW Conference, Marseille, June 2023*. Vol. 138. Proceedings e Report. Florence : Firenze University Press. <https://doi.org/10.36253/979-12-215-0413-2>.
- « ISO 28500:2017 Information and Documentation – WARC File Format ». 2017. <https://www.iso.org/standard/68004.html>.

- Lavigne, Michel. 2005. « Regard rétrospectif sur les CD-ROM culturels ». *Entrelacs*, décembre. École nationale supérieure d'audiovisuel (Toulouse). <https://hal.science/hal-03181146>.
- Magro, Sébastien, et Omer Pesquer. 2016. « Une archéologie des premiers sites web de musées en France ». <https://sebastienmagro.net/2016/03/14/une-archeologie-des-premiers-sites-web-de-musees-en-france/>.
- Pédauque, Roger T. 2006. « Document et modernités ». https://archivesic.ccsd.cnrs.fr/sic_00001741.
- Segalen, Martine. 2019. « Le Musée national des Arts et Traditions populaires, 1936-2005. Récit d'un brillant fiasco. Première partie : Une si longue naissance (1880-1980) ». *Encyclopédie Bérose des histoires de l'anthropologie*. Paris : CNRS - Ministère de la Culture. <https://www.berose.fr/article1685.html>.
- Vidal, Geneviève. 2018. « Médiation muséale en ligne ». In *La médiation numérique muséale. Un renouvellement de la diffusion culturelle*, 17-40. Labyrinthes. Pessac : Presses Universitaires de Bordeaux. <https://doi.org/10.4000/books.pub.32558>.