

Collège de France



Les archives de fouilles : modes d'emploi | Sandra Zanella, Jean-Pierre Brun, Martine Denoyelle, et al.

Missions et archives de fouille. Entre la production et la conservation

Alexandre Rabot

<https://orcid.org/0000-0001-5425-918X>

Texte intégral

- 1 La question des données de la recherche est, depuis quelques années, au cœur d'un vif débat dans le domaine des Sciences humaines et sociales, et l'archéologie, grande productrice de données, en volume et en variété typologique, est particulièrement concernée.
- 2 Les défis sont nombreux et n'ont commencé à être relevés que tardivement. Certaines évolutions déterminantes ont eu lieu sans que les cadres théoriques et pratiques indispensables n'aient été réellement posés. On pensera par exemple au passage au numérique : toute la documentation archéologique l'est désormais, nativement ou *a posteriori*.
- 3 La question sera ici abordée à partir du cadre, du module qu'est une mission archéologique, de taille petite ou moyenne, liée à un laboratoire (type Unité mixte de recherche) sans équipement spécifique propre.
- 4 La mission est en même temps un lieu de production (données immédiates) et un lieu de conservation (données anciennes, produites par la Mission elle-même ou par les d'organisations qui l'ont devancée).
- 5 Elle occupe une position intermédiaire entre le chercheur seul et l'institution, entre le producteur de la donnée et celui qui en assurera, à très long terme, la pérennité.
- 6 Cette intervention ne propose rien d'exemplaire mais veut, modestement, dans une logique de partage de 'bonnes pratiques' présenter les choix et les outils qui ont pu être mis en place dans la limite des moyens alors à disposition.

Constats sur les environnements

- 7 Les équipes des missions archéologiques sont très composites, quant aux spécialités scientifiques (en constante diversification), au rattachement (institutions nombreuses et variées voire, parfois, *a contrario*, pas encore d'insertion pérenne), à la résidence physique. Il y a peu ou pas d'espace commun collectif. L'équipe n'est réunie souvent que le temps du terrain (et pas toujours de façon synchrone). Cette configuration n'est pas sans effet « centrifuge », dispersif, et ce malgré les outils de

Le site utilise des cookies et collecte des informations personnelles vous concernant.

Pour plus de précisions, nous vous invitons à consulter notre **politique de confidentialité** (mise à jour le 25 juin 2018).

En poursuivant votre navigation, vous acceptez l'utilisation des cookies.

Fermer

commun (en particulier sur les questions de création et traitement des données).

- 9 À l'échelle des laboratoires, il n'y a encore que peu de stratégie collective, d'outils repérables, partagés, validés et largement adoptés. Il n'y a pas d'agent ou de service dédié, guère de « bonnes pratiques » unanimement reconnues, peu de plates formes communes.
- 10 Par ailleurs, la relation des chercheurs à leurs données est encore très « personnelle » ou « exclusive » (les guillemets s'imposent tant cette question est vaste, et humainement et scientifiquement délicate). Ils ont été souvent les seuls porteurs et supports de ses données, sans bénéficier de relais efficace ou de partenaire face à cette charge, face aux risques inhérents aux données (perte physique, perte de lisibilité, d'intelligibilité...).
- 11 En outre, la distinction « dynamique », « positive », entre la démarche « publication » (la modalité historique de partage, la plus légitime, la plus valorisée) et la démarche « diffusion de la donnée » est encore souvent à construire. La première semble obligatoirement contenir la seconde (c'est-à-dire pas de diffusion de données sans publication). Cela favorise des mécanismes de « rétention » qui préparent mal à poser les questions de fond sur les données elles-mêmes.
- 12 Enfin, l'adoption d'outils et de pratiques peut être ralentie par d'autres facteurs humains, comme l'environnement intellectuel (le milieu SHS est parfois moins disponible à des outils techniques, hors spécialités scientifiques, qui sont peu présentés et enseignés durant les parcours de formation), ou la « pyramide des âges » (la plupart des chercheurs pouvant être au seuil de la retraite, donc moins disponibles à l'innovation « pratique »).

Constats généraux sur l'état de la prise en charge des données

- 13 La pérennité des données n'est pas systématiquement assurée *via* des outils ou des plates formes sécurisées. Elles se retrouvent rarement sur des supports communs (elles sont souvent dispersées entre divers « disques durs », fixes ou mobiles). Elles ne « confluent » pas.
- 14 Les supports communs ne sont pas forcément institutionnels. On n'en maîtrise pas vraiment les limites, les logiques d'exploitation (souvent purement commerciales d'ailleurs, par ex. DropBox, WeTransfer).
- 15 Le nombre de données augmente de façon exponentielle ; elles sont souvent simples à produire et leur coût de réalisation est faible voire nul).
- 16 L'adoption de nouveaux logiciels peut en accroître le nombre. On pensera en particulier aux travaux de photogrammétrie et d'imagerie 3D ou de SIG.
- 17 Les types de formats se diversifient grandement. Si certains sont très diffusés, libres ou ouverts, d'autres sont plus rares, verrouillés ou éphémères (sous licence, dépendants de logiciels peu diffusés ou très spécifiques).
- 18 Les données sont souvent « muettes ». Leur dénomination est aléatoire (par ex. les noms de fichiers sont ceux produits automatiquement, par les appareils photo non paramétrés). Aucune donnée qualitative n'y est intrinsèquement associée (pas de données embarquées). Les plans de classement sont souvent induits, pas ou peu remis à plat, et élaborés sans recours à des référentiels.

Objectifs

- 19 Suite aux expériences et aux réflexions collectives au sein du laboratoire¹, il est possible de poser certains cadres et quelques objectifs...
- 20 De façon générale, il est essentiel d'être conscient que l'on se trouve au tout début d'une chaîne documentaire qui se poursuivra bien au-delà de nous et que nos « bonnes pratiques », aujourd'hui, favoriseront la pérennité et la transmission de nos données, demain et après-demain...
- 21 Il semble important de s'appuyer au maximum sur des solutions largement utilisées, validées, sans créer forcément de solutions « internes », trop dépendantes de moyens locaux, souvent ponctuels ou temporaires (personnel, outil).
- 22 Quatre axes peuvent être dessinés : A) l'information des données ; B) le classement ; C) le choix des formats ; D) les supports de partage et de conservation.

A. Information des données

- 23 Cette étape peut être menée grâce à des logiciels très accessibles, gratuits (comme Xn View²) ou payants (comme Adobe Bridge³).

Ce site utilise des cookies et collecte des informations personnelles vous concernant.

Pour plus de précisions, nous vous invitons à consulter notre [politique de confidentialité](#) (mise à jour le 25 juin 2018).

En poursuivant votre navigation, vous acceptez l'utilisation des cookies.

Fermer

l'expérience montre qu'il est important qu'il garde du sens pour l'utilisateur et qu'il puisse produire, par simple affichage, son propre classement logique. La syntaxe doit être simple et intelligible. Elle doit nécessiter le moins de concertation possible au cours du travail et éliminer la possibilité de doublon. Les préconisations du SRA d'Alsace⁴ nous ont permis, par exemple, d'établir une séquence sur une suite d'information (auteur – nom de mission – année de mission – type de document – incrémentation – extension, ou, – jour de création – incrémentation – extension), par ex. AR_SMT_2016_NIMFO_0090.jpg, ou AR_SMT_2016_NIMFO_160125_0001.jpg.

2. Renseignements systématiques de certaines métadonnées

- 25 Ces métadonnées⁵ peuvent être embarquées dans le document ou dans un document annexe, *via* des champs normalisés. Elles sont essentielles pour retrouver la donnée grâce aux informations dont elle est porteuse.
- 26 Nous avons préféré, pour des raisons ergonomiques et de stabilité du lien document-métadonnée, installer l'information dans le document lui-même en utilisant le protocole IPTC⁶. Ce système, très généraliste, imaginé autrefois pour la presse et les médias, présente des avantages (très grande présence, usage et diffusion) et des limites (champs généralistes, parfois peu adaptés à des données scientifiques spécifiques). Ces dernières limites sont en partie dépassées par la version IPTC Plus et très récemment l'IPTC Cultural Heritage Panel⁷. D'autres systèmes sont également disponibles et activables sur les logiciels courants (par ex. le VRA Core, anglophone⁸).
- 27 Il est possible d'exporter les données embarquées vers des tableurs type .csv (grâce par ex. à l'outil IPTC-Plus Metadata).
- 28 Nous avons privilégié l'information de certaines métadonnées, jugées essentielles (elles peuvent être attribuées par lot plus ou moins large, en fonction de leur caractère commun à tout ou partie des données) :
- Auteur (champ *creator*), et certaines informations afférentes (fonction de l'auteur, adresse, ville, état/province, code postal, pays) ;
 - Localisation (champ *location*) : lieu, ville, état/province, pays, code de pays ISO ;
 - Description (champ *description*), en texte libre ;
 - Droits (champ *copyright*) : personne physique, organisation ou institution associé et état des droits (domaine public, soumis à des droits, inconnu) ; la notion de copyright est approximative et par défaut dans la mesure où elle ne relève pas du droit français⁹ ;
 - Mots-clefs (champ *keywords*). On distinguera les « identifiants », ceux en usage dans un projet ou une équipe donnée (par ex. numéros d'US, de structures, d'objets, d'inventaire) et « les mots-clefs », pour lesquels on tentera d'utiliser un référentiel extérieur (par ex. le thésaurus Pactol du réseau Frantiq, *via* l'OpenTheso¹⁰).
- 29 D'autres champs, « secondaires », peuvent être exploités, dont celui gérant la géolocalisation (champ *GPS*) quand il y a, par exemple, un enjeu de visualisation cartographique (couverture de prospection etc.).

B. Classement

- 30 Le plan de classement doit articuler deux logiques : reprendre des structurations réfléchies et éprouvées par ailleurs¹¹ et développer des arborescences propres à chaque projet.

C. Formats

- 31 La pérennité et l'interopérabilité des formats conditionnent l'exploitation à long terme de nos données. On privilégiera donc ceux qui sont les plus standards, libres et ouverts. On n'hésitera donc pas à stabiliser un document *via* un export vers un format adapté (par exemple .csv pour un tableur, .svg pour un document vectoriel, .obj pour un modèle 3D). Les préconisations du Cinès sont précieuses en la matière¹².

D. Les supports de partage et de conservation

- 32 Cette démarche a un objectif pratique mais aussi fortement symbolique : « mettre en commun » et faire de la donnée un « bien commun », en maintenant cependant, sans ambiguïté, le lien entre celle-ci et son auteur, son contenu, son contexte.

Le site utilise des cookies et collecte des informations personnelles vous concernant.

Pour plus de précisions, nous vous invitons à consulter notre [politique de confidentialité](#) (mise à jour le 25 juin 2018).

En poursuivant votre navigation, vous acceptez l'utilisation des cookies.

Fermer

- des outils internes (par ex. solution type OwnCloud¹⁴) ; la donnée est regroupée sur un poste ; elle est sauvegardée automatiquement sur un serveur et les mises à jour sont synchronisées ; elle est accessible *via* une interface web ;
 - des solutions de type GED (Gestion Électronique de Document, par ex. *via* Nuxeo¹⁵), avec une logique de gestion de document, et pas simplement de conservation ; le document a un parcours, une évolution qui peuvent être suivis, validés, partagés ;
 - des serveurs externes (par ex. solution type Sharedoc de Huma-Num¹⁶) ; les documents sont déposés sur un serveur et accessible suivant un jeu d'autorisation (lié à un annuaire) ou de liens web partagés.
- 35 Tous ces choix gagneront à être explicités, synthétisés et décrits dans le « plan de gestion de données » (*data management plan*, DMP). Celui-ci, simple ou très articulé en fonction de l'avancement et la complexité documentaire, est une démarche payante pour les données elles-mêmes et valorisante pour les projets qu'elles accompagnent. Des guides de bonnes pratiques sont déjà disponibles ainsi que des outils en ligne (par ex. DMP on line¹⁷).
- 36 Au-delà de leurs aspects apparemment « techniques », ces questions sous-tendent et accompagnent les évolutions en cours touchant aux activités et stratégies du chercheur contemporain. Son rôle de producteur et gestionnaire de données renforce et assoie ses fonctions plus « purement » scientifiques. À terme, on appréhendera, voire évaluera, un chercheur, des équipes, des projets non plus seulement à travers leurs productions classiques, leurs publications, mais également *via* les données et ensemble de données qu'ils auront su créer, faire vivre, partager et transmettre...

Notes

1. L'unité mixte bénéficie de la présence d'une Ingénieure d'étude CNRS, particulièrement active et réactive sur ces questions (cf. <http://www.hisoma.mom.fr/annuaire/morlock-emmanuelle>).
2. <http://www.xnview.com/fr>.
3. <http://www.adobe.com/fr/products/bridge.html>.
4. <https://docpatdrac.hypotheses.org>.
5. Pour une définition, cf. par ex. http://www.adbs.fr/metadonnees-17808.htm?RH=OUTILS_VOC.
6. <https://iptc.org>.
7. <https://iptc.org/news/culturalheritagepanel>.
8. <http://core.vraweb.org>.
9. Cf. le tableau comparatif proposé par la Direction des affaires juridiques du CNRS, <https://www.dgdr.cnrs.fr/daj/propriete/droits/Docs/comparatif-auteur-copyright.pdf>.
10. <http://frantiq.mom.fr/fr/thesaurus> et <http://pactols.frantiq.fr/opentheso>.
11. Cf. le travail précieux déjà accompli à Nanterre par l'équipe des archives d'E. Bellon (Guide d'archivage, p. 6-11, cf. <http://www.mae.u-paris10.fr/service-des-archives-de-la-mae>).
12. <https://facile.cines.fr>.
13. Le choix ou la mise en place de ces outils ont bénéficié de l'aide des ingénieurs du service PSIR de la MOM (cf. <http://www.mom.fr/unite-de-service-et-de-recherche/pole-systeme-d-information-et-reseaux/presentation>).
14. <https://owncloud.org>.
15. <https://www.nuxeo.com/fr>.
16. <http://www.huma-num.fr>.
17. <https://dmponline.dcc.ac.uk>.

Auteur

Alexandre Rabot

**Ingénieur d'étude en Archéologie, laboratoire HiSoMA,
UMR 5189, Maison de l'Orient et de la Méditerranée**

Du même auteur

Kition-Bamboula VII. MOM Éditions. 2018

Ce site utilise des cookies et collecte des informations personnelles vous concernant.

Pour plus de précisions, nous vous invitons à consulter notre **politique de confidentialité** (mise à jour le 25 juin 2018).

En poursuivant votre navigation, vous acceptez l'utilisation des cookies.

Fermer

Indices (fouilles françaises dans la nécropole de Kition-Pervolia) in Kition-Bamboula VII, MOM Éditions, 2018

Tous les textes

© Collège de France, 2017

Conditions d'utilisation : <http://www.openedition.org/6540>

Référence électronique du chapitre

RABOT, Alexandre. *Missions et archives de fouille. Entre la production et la conservation* In : *Les archives de fouilles : modes d'emploi* [en ligne]. Paris : Collège de France, 2017 (généré le 01 novembre 2019). Disponible sur Internet : <<http://books.openedition.org/cdf/4909>>. ISBN : 9782722604650. DOI : 10.4000/books.cdf.4909.

Référence électronique du livre

ZANELLA, Sandra (dir.) ; et al. *Les archives de fouilles : modes d'emploi*. Nouvelle édition [en ligne]. Paris : Collège de France, 2017 (généré le 01 novembre 2019). Disponible sur Internet : <<http://books.openedition.org/cdf/4859>>. ISBN : 9782722604650. DOI : 10.4000/books.cdf.4859.

Compatible avec Zotero

Ce site utilise des cookies et collecte des informations personnelles vous concernant.

Pour plus de précisions, nous vous invitons à consulter notre **politique de confidentialité** (mise à jour le 25 juin 2018).

En poursuivant votre navigation, vous acceptez l'utilisation des cookies.

Fermer