

RECOMMANDATIONS DU CONSORTIUM 3D SHS

PÉRENNISATION DES DONNÉES 3D

L'ARCHIVAGE PÉRENNE DES MODÈLES NUMÉRIQUES 3D POUR LES SHS

ÉDITION 2019



Bruno Dutailly (CNRS, Archéovision - UMS 3657)

Sylvie Eusèbe (INRAP)

Valentin Grimaud (CNRS, CREAAH, LARA UMR 6566)

Nicolas Lefèvre (UFR HSS - CIREVE)

Matthieu Quantin (ECN, Nantes)

Sarah Tournon-Valiente (CNRS, Archéovision - UMS 3657)

1. CONTEXTE DE TRAVAIL

1.1. Objectifs

Depuis quelques décennies, les représentations tridimensionnelles des vestiges archéologiques sont utilisées pour aider à la compréhension, la conservation et la médiation. Avec l'accès à des ordinateurs de plus en plus puissants et la diffusion des savoir-faire, l'usage des modèles 3D s'est largement et rapidement démocratisé. Ces modèles tiennent une place particulière dans la connaissance de l'objet archéologique étudié.

Cependant, il n'existe à ce jour aucune solution pérenne de sauvegarde de ces données. Or, de plus en plus de modèles sont produits chaque année dans le cadre de recherches. Dans le meilleur des cas, ils sont sauvegardés sur les machines des chercheurs avec éventuellement un stockage supplémentaire, considéré comme plus sécurisé, sur un autre poste ou un autre support. À long terme, ces conditions ne sont pas satisfaisantes pour conserver ces données.

Une réflexion commune menée par la *TGIR* Huma-Num⁵, le *CINES*⁶ et le laboratoire Archéovision⁷ (Bordeaux) – formalisée à travers le *Consortium 3D SHS*⁸ – a permis d'aboutir à la création d'un service d'archivage : de la même façon qu'un texte, une image ou un document sonore peuvent être archivés, l'objectif est de rendre possible l'archivage des modèles 3D.

Ce document est la formalisation d'une étape de la réflexion. Il rend compte des décisions qui ont été prises tout en ayant la volonté d'être un guide pour les producteurs de 3D qui souhaiteraient archiver leurs modèles. La

procédure d'archivage au *CINES* sera progressivement détaillée afin de guider le producteur dans la constitution d'un lot d'archives.

Même si tous les modèles n'ont pas vocation à être archivés au *CINES*, la méthodologie et les préconisations sont valables pour tout acteur de l'archivage, que ce soit au sein d'un laboratoire, d'une université, ou d'une autre structure.

1.2. Nécessité d'archiver la 3D en archéologie

Quels que soient leurs objectifs (étude, conservation, diffusion – comme le montre l'illustration ci-dessous), le modèle 3D et ses documents associés véhiculent des connaissances scientifiques et techniques, aussi bien passées que nouvelles. Par conséquent, il est légitime de vouloir les transmettre et donc de les archiver au même titre que les autres vecteurs de savoirs archéologiques (texte, photo, dessin, empreinte, base de données, etc.). Comme pour tous ces vecteurs, documenter la démarche qui a conduit au modèle 3D contribue à sa valeur scientifique : décrire le cheminement intellectuel suivi, ses objectifs, ses limites, ses méthodes, ses degrés de certitude et ses sources.

⁵ Très Grande Infrastructure de Recherche Huma-Num. Cf. ([site](#) consulté le 24/07/2019)

⁶ Centre Informatique National de l'Enseignement Supérieur. Cf. ([site](#) consulté le 24/07/2019).

⁷ Cf. ([site](#) consulté le 24/07/2019).

⁸ Cf. ([site](#) consulté le 24/07/2019).

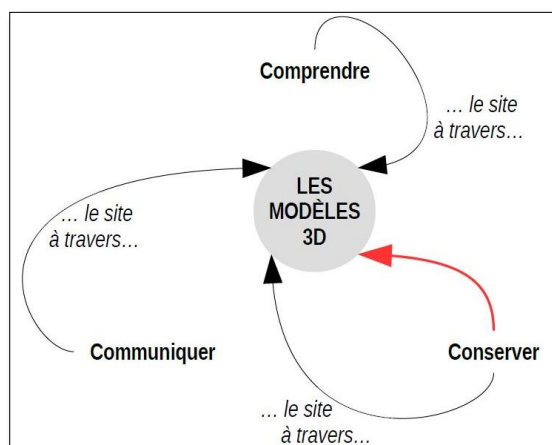


Figure 1 - Les modèles 3D sont au cœur de trois enjeux et la conservation tient un rôle particulier. Le présent travail ne s'attache qu'à la relation matérialisée par la flèche rouge

Satisfaire à la rigueur de la démarche scientifique conduit donc à porter une grande attention aux métadonnées : les documents archivés aujourd'hui seront les sources de demain. Il est du devoir du chercheur de rendre ces futures sources aussi réappropriables que possible, en réfléchissant à ce dont ses successeurs pourraient avoir besoin, au regard des informations qui lui sont utiles à propos des sources auxquelles il fait appel aujourd'hui et de celles dont il regrette l'absence.

La modélisation 3D en archéologie concerne les vestiges mobiliers et immobiliers, quelles que soient leur taille, leur morphologie, leur localisation géographique et leur époque. On distingue trois catégories de modèle 3D : le modèle **tel que saisi**, le modèle **tel que construit** et la **reconstitution**. Le premier se veut être une image la plus exacte possible de l'état actuel du vestige, le second apporte une description sémantique en segmentant le maillage objet par objet, alors que la reconstitution vise à présenter un état antérieur, au moment de sa construction et/ou de toute étape intermédiaire jugée pertinente.

Le modèle est un outil permettant d'accéder à une compréhension du vestige complémentaire ou différente de celle offerte par les autres outils

de visualisation, son utilisation participe de différentes façons à l'étude.

En documentant l'existant, le modèle de terrain sert de support visuel pour l'analyse, l'aide au relevé, ainsi qu'à la présentation des résultats. Il contribue à l'observation (point de vue différent de celui dans la réalité), à la description (dont la possibilité de mesurer et d'éditer des ortho-images mosaïques) et à la compréhension (organisation, structure, fonction, etc.).

Le modèle de restitution permet de s'interroger sur les parties manquantes de l'objet réel étudié. Comme la maquette de restitution réelle, ce modèle 3D permet de s'interroger sur la morphologie, l'agencement et la couverture des volumes, ainsi que la circulation, les ouvertures et l'évolution chronologique des vestiges. La simulation de phénomènes physiques, difficilement accessible auparavant, est rendue plus aisée par cette modélisation 3D. Enfin, d'un point de vue méthodologique, la compréhension immédiate du modèle favorise le travail collaboratif entre chercheurs de disciplines jusqu'à présent peu en interaction.

La visualisation 3D apparaît plus intuitive que les traditionnels plans et coupes (du moins aujourd'hui et dans la culture occidentale). La possibilité de consulter le vestige pendant une durée plus longue que celle de la fouille, dans un autre contexte (à tout moment et en tout lieu), favorise l'émergence de nouveaux questionnements, de nouvelles hypothèses. Les attentes heuristiques liées à la modélisation 3D sont prometteuses et renforcent la nécessité de son archivage.

1.3. Contexte

La gestion des archives dans le domaine du patrimoine culturel – et plus largement dans le domaine des sciences humaines et sociales – a engendré assez peu d'études : les initiatives résultent dans un premier temps de politiques nationales avant de devenir un sujet porté à une échelle transnationale.

Lignes directrices

En 2003, l'UNESCO se saisit de la question à travers la *Charte sur la conservation du patrimoine numérique*⁹. Cela ne concerne pas directement les représentations virtuelles en archéologie, mais la menace de la perte (*article 3*), la nécessité d'agir (*article 4*) et l'aspect pérenne de l'information numérique (*article 5*) sont formellement établis. Ensuite, une série de mesures requises est annoncée : il s'agit d'élaborer des stratégies et des politiques (*article 6*), de sélectionner ce qu'il convient de conserver (*article 7*), de protéger le patrimoine numérique (*article 8*) et de préserver le patrimoine culturel (*article 9*). Enfin, les responsabilités (*article 10*), les partenariats et les coopérations (*article 11*) et le rôle de l'UNESCO (*article 12*) sont définis.

En 2009, une charte est rédigée plus spécifiquement pour la numérisation du patrimoine culturel : la *Charte de Londres* (Denard 2009)¹⁰. Elle concerne néanmoins tous les types de documents : images, audiovisuel, textes, etc. Quelques années plus tard, après l'explosion des acquisitions 3D due à la démocratisation de la photogrammétrie, l'usage de la 3D sera précisé dans la *charte de Séville*¹¹ (2011) à travers huit principes : l'interdisciplinarité, la définition des objectifs, la complémentarité avec les moyens existants, l'authenticité, la rigueur historique, l'efficacité, la transparence scientifique et la formation / évaluation. Cependant, si le terme de base de données est largement utilisé, les questions plus spécifiques liées à l'archivage ne sont pas abordées.

Le Conseil européen de l'archéologie diffuse en 2014 un document synthétisant les bonnes pratiques pour l'archivage des données (Perrin et al. 2014). Il fournit ainsi un ensemble de

recommandations, tant du point de vue de la gestion des données numériques que de celui des objets découverts au cours de fouilles. Pour cela, une première partie s'attache à rappeler les principes essentiels et les normes qui en découlent. Des exemples concrets sont ensuite proposés pour illustrer le propos et faciliter la mise en œuvre des préconisations. Une dernière partie expose des exemples de coordination entre les différents acteurs du « projet archéologique », ainsi qu'un phasage des opérations afin d'assurer le bon déroulement des travaux.

Ces documents internationaux n'abordent finalement pas la question de l'archivage, même si ce thème commence à apparaître dans les dernières recommandations. Le sujet est en revanche porté de manière plus ferme par des initiatives nationales.

Initiatives nationales opérationnelles

Depuis 1996, la question de la sauvegarde des données numériques en archéologie est traitée par *Archaeology Data Service*. Ce service est né d'un consortium regroupant huit départements d'archéologie des universités du Royaume-Uni¹². L'objectif est de fournir des conseils sur la façon de créer et de gérer l'ensemble des données numériques. Le schéma MIDAS¹³ est publié dès 1998 (dernière mise à jour en 2012) afin de spécifier et classer les informations nécessaires à la documentation du patrimoine culturel. Depuis 2011, la version 3 d'un guide de bonnes pratiques est publiée sur le site Internet du projet¹⁴ et concerne les systèmes d'information géographique, les données 3D (télédétection), la prospection géophysique, la rédaction de rapports de fouilles, ou encore l'émergence de la réalité virtuelle. Le service joue désormais un rôle central dans l'archivage des

⁹ Cf. ([site](#) consulté le 24/07/2019).

¹⁰ Cf. ([site](#) consulté le 24/07/2019).

¹¹ Cf. ([site](#) consulté le 24/07/2019).

¹² Universités de Birmingham, de Bradford, de Glasgow, de Kent à Canterbury, de Leicester, d'Oxford, de Newcastle et de York, en partenariat avec le Council for British Archaeology.

¹³ Cf. ([site](#) consulté le 24/07/2019).

¹⁴ Cf. ([site](#) consulté le 24/07/2019).

résultats de fouilles au Royaume-Uni en recueillant les rapports des opérations financées par l'administration publique et des publications scientifiques comme la revue en ligne *Internet Archaeology* ou la *Society of Antiquaries of Scotland*. À ce titre, un guide pratique pour les utilisateurs est disponible (ADS 2014). Toute la documentation réunie a pour ambition d'être mise à disposition du plus grand nombre et donc d'être consultée. La procédure opérationnelle concernant les données 3D est absente (voire contradictoire). L'institution procède au cas par cas. Elle propose néanmoins un outil en ligne (*ADS easy*) pour le chargement des données à archiver, non-compatible avec les données 3D. La visualisation des modèles 3D en ligne étant alors un problème, un projet européen est mis en place afin de le résoudre : c'est dans ce contexte que le visualiseur *3DHop* est créé¹⁵.

En 2005, l'Université de Virginie (USA) lance le service « *Serving and Archiving Virtual Environments* »¹⁶ (SAVE) afin de diffuser, archiver et publier les modèles 3D produits en archéologie avec les standards des publications scientifiques, c'est-à-dire avec un comité de lecture. Cependant, la documentation disponible est très limitée et l'absence de mise à jour du site Internet depuis 2009 indique que le projet ne semble pas avoir rencontré l'audience escomptée.

Un an plus tard, en 2006, l'université d'État d'Arizona crée le « *Digital Archaeological Record* » (*tDAR*)¹⁷ grâce à un financement de la *National Science Foundation*. L'objectif est de mettre à disposition des chercheurs un ensemble d'outils permettant de normaliser les données produites pour faciliter leur partage et leur conservation. Cela se traduit par la création d'un dépôt numérique de production – à l'échelle nationale dans un premier temps. Un rapprochement avec *Archaeology Data Service*, dès 2008, conduit à la fondation de *TAG* (*Transatlantic Archaeology*

Gateway). Le but de cette organisation est de fournir la même qualité de prestation en termes d'archivage à l'échelle internationale que ce qui est proposé au niveau national. En effet, *ADS* permet aux archéologues britanniques d'archiver leurs données et de les rendre disponibles sur le long terme ; *tDAR* propose la même chose pour les scientifiques américains. Cela est possible grâce à une harmonisation des recommandations et des structures de données en s'appuyant principalement sur le schéma *MIDAS*. La procédure opérationnelle concernant les données 3D est absente, l'institution procède toujours au cas par cas. Le guide publié en 2014 par *ADS*¹⁸ est ainsi ponctué de *nota bene* indiquant :

“if there is a special reason for needing data sets accessioned more quickly, please contact the ADS in advance to discuss your requirements “(page 8).

“if your data type is not here please contact ADS (help@archaeologydataservice.ac.uk) to discuss if we can archive your data“ (page 20).

“if the collection you are depositing is a large collection of text documents only, such as the results of a journal digitization project, please contact ADS directly as we may also require additional documentation “(page 22).

La photogrammétrie fait exception, avec des métadonnées très complètes, mais les exigences sont lourdes, et semblent déconnectées des pratiques en SHS¹⁹.

Ce schéma ne sert pas uniquement à documenter et archiver les données produites pendant l'étude du patrimoine culturel ; il est également utilisé pour diffuser les contenus sur l'Internet. Il fait ainsi partie du schéma *CARARE*²⁰, produit par le consortium européen *3D-ICONS*, dont l'unique objectif est d'alimenter la médiathèque virtuelle *Europeana*.

¹⁵ Cf. ([site](#) consulté le 24/07/2019).

¹⁶ cf. ([site](#) consulté le 24/07/2019)

¹⁷ cf. ([site](#) consulté le 24/07/2019)

¹⁸ cf. ([site](#) consulté le 24/07/2019)

¹⁹ cf. ([site](#) consulté le 24/07/2019)

²⁰ cf. ([site](#) consulté le 24/07/2019)

Enfin, une dernière initiative doit être mentionnée : le projet *Archéogrid*²¹ porté par Archéovision dès 2007. Il s'agit de permettre le dépôt de fichiers sur un serveur avec la possibilité de les consulter. Afin de les retrouver, un ensemble de métadonnées est renseigné. Devant l'ambition du programme, la volonté de développer la réflexion sur l'archivage à l'échelon national conduit à créer en 2014 le Consortium 3D SHS²² pour résoudre notamment cette question. La philosophie vise à bénéficier de l'expertise du CINES sur cette thématique – les tenants et aboutissants seront détaillés plus loin dans ce document.

Il faut enfin noter que des initiatives ont émergé du milieu des archivistes, notamment sous l'égide du projet européen *E-ARK « European Archival Records and Knowledge Preservation »* mené de 2014 à 2017. À l'issue du programme, un logiciel a été mis à la disposition de la communauté pour générer des archives : *RODA-in*²³.

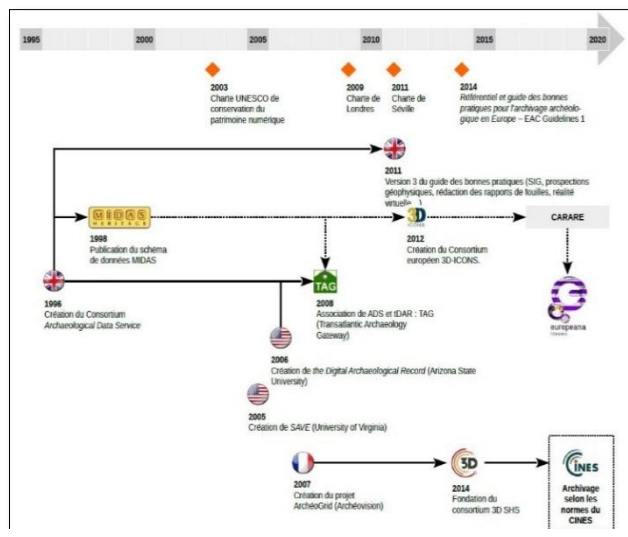


Figure 2 - Schéma retraçant des projets concernant l'archivage des données produites pendant l'étude du patrimoine culturel.

Plan de gestion de données

La réflexion présentée dans ce document entre également en résonance avec la politique

européenne préconisant l'ouverture des données scientifiques à travers le programme *H2020*²⁴. Cela se traduit par la rédaction de Plans de Gestion des Données (PGD – ou *DMP* en anglais pour *Data Management Plan*).



Le thème est très vaste et stimule aujourd'hui de nombreuses réflexions – ainsi ne ferons-nous que l'effleurer. Pour approfondir le sujet, se reporter à « Une introduction à la gestion et au partage des données de la recherche »²⁵ (Cartier et al. 2015)

Pour résumer, le PGD couvre les champs essentiels permettant d'anticiper le projet d'archivage et la dissémination des données. Le travail mené par les universités Paris-Diderot, Sorbonne-Paris-Cité et Paris Descartes (Cartier et al. 2015) propose un modèle de plan de gestion de données :

En-tête :

- **Section 1** : Information sur le projet
- **Section 2** : Responsabilité des données
- **Section 3** : Ressources nécessaires à la mise en œuvre du PGD
- **Section 4** : Jeux de données
 - **Section 4.1** : Description du jeu de données
 - **Section 4.2** : Au cours du projet – Stockage, accès et sécurité des données
 - **Section 4.3** : Métadonnées : documentation et organisation des données
 - **Section 4.4** : À l'issue du projet : dissémination des jeux de données
- **Section 5** : Sélection et archivage

La création d'un tel document en début de projet permet d'obtenir une bonne visibilité de la gestion des documents et sa meilleure mise en œuvre. En raison de la multiplicité des savoir-faire requis pour couvrir tous les aspects du projet, il est

²¹ cf. ([site](#) consulté le 24/07/2019)

²² Cf. ([site](#) consulté le 24/07/2019)

²³ cf. ([site](#) consulté le 24/07/2019)

²⁴ Cf. ([site](#) consulté le 24/07/2019).

²⁵ Cf. ([site](#) consulté le 24/07/2019).

peu rationnel de confier l'ensemble des tâches (de l'acquisition à l'archivage en passant par le traitement des modèles) à une seule personne ; la définition du rôle de chacun est primordiale et les compétences comme les outils sont spécifiques à chacune des étapes.

La réflexion portée par l'atelier Archivage du Consortium 3D SHS permet de documenter la partie « 4.3 : métadonnées : documentation et organisation des données » de la rédaction d'un plan de gestion des données pour un projet s'appuyant sur des modèles 3D (en archéologie...) et de faciliter la mise en œuvre de la section « 5 : sélection et archivage ».

1.4. Les métadonnées

Le but ici n'est pas de créer un nouveau « guide de bonnes pratiques » ou un état de l'art sur les métadonnées pour le patrimoine et l'archéologie. Il s'agit de synthétiser les fondements pour le lecteur non averti.



Étymologiquement, les métadonnées sont des « données de données ». Il s'agit d'un ensemble d'informations plutôt courtes.



*Elles peuvent être contenues dans le fichier qu'elles renseignent, à l'instar des données exif d'une photo. Elles peuvent également être placées dans un fichier séparé, généralement au format *.xml.*

Tour d'horizon des schémas de métadonnées

On considère trois types de métadonnées :

Contenus (contents) : ils indiquent ce qu'il faut renseigner pour décrire ses données, c'est un guide avec des indications théoriques. Il s'agit de bonnes pratiques. Le niveau d'abstraction est élevé, il n'est pas exploitable directement pour décrire ses données, mais fait autorité pour construire un schéma de métadonnées (de type structure). Ils ne nous intéressent pas directement dans notre cas d'archivage. Ils sont rares et

suffisants. Ex : *CCO, DublinCore Guidelines, ADS Guide*.

Valeurs (values) : ce sont des notices d'autorité, des thésaurus, des vocabulaires contrôlés. Ces schémas sont créés par des institutions influentes au niveau national/mondial (musées nationaux, ministères). Il s'agit d'un système de mots réservés et organisés. Ces schémas peuvent être instanciés. Ils définissent ainsi comment se référer à une œuvre d'art (La Joconde ou Mona Lisa ?) ou à un artiste. Ils peuvent aussi définir des termes techniques et leur usage. Certains sont multilingues. Les éléments reconnus sont rares et suffisants, car ils doivent faire autorité. Ex : *CONA, ATT*, les vocabulaires du Ministère de la Culture. Ils sont souvent complétés par des vocabulaires contrôlés « locaux » en fonction d'un besoin (par exemple un musée de l'automobile aura besoin d'une terminologie plus aboutie en mécanique). Des projets comme *PACTOLS*²⁶ en archéologie unifient les efforts de création de thésaurus dans une discipline.

Structure (structure) : elles indiquent ce qu'il faut renseigner sous forme de « questions » explicites (pour un livre par exemple le titre, ou la date d'édition). Elles sont souvent spécialisées dans un domaine (auteurs, peinture, archéologie...). Certains schémas sont très génériques et sont utilisés dans de nombreux domaines. Les usages définissent également des clivages (archivage, conservation, diffusion). Ils sont nombreux, car spécialisés dans différents domaines.

Deux grandes technologies de structures s'opposent :

Les descripteurs plats ou faiblement structurés avec une grande liberté dans les valeurs à renseigner (ex *Dublin-Core, CDWA Lite, CINES-PAC*). La qualité des données décrites limite le partage et la diffusion, mais avec une grande simplicité, cela permet la gestion, la valorisation et la pérennité des données pour une institution. Ces

²⁶ [site](#) consulté le 24/05/2019

descripteurs sont faiblement interopérables (plus ou moins étanche les uns par rapport aux autres).

Les ontologies (de domaine) : organisation en entités-relations destinée au partage ou à la formalisation de données. Les deux grands objectifs sont (1) de formaliser les connaissances d'un domaine pour les stocker explicitement, et (2) de pouvoir « raisonner automatiquement » sur ces données pour en produire de nouvelles. Dans le cas du web sémantique, les données sont réticulées, c'est-à-dire qu'une unique et même ressource est appelée (liée) à chaque usage. Cela rajoute un troisième objectif : (3) Partager les données, utiliser Internet comme une immense base de données distribuée. Ex CIDOC-CRM, Foaf, Bibframe

Dans les 2 cas, les données décrites sont réutilisables au cas par cas (faiblement interopérables), à moins que des vocabulaires contrôlés, thésaurus ou notices d'autorité aient été utilisés pour les décrire.

Un tableau en annexe traite un grand nombre de schémas de métadonnées couramment rencontrés.

 [cf. annexe](#) – Schémas de données couramment utilisées pour la documentation du patrimoine culturel.

Nous proposons un schéma de métadonnées de type « structure » non-ontologique, compatible avec le schéma de structure du *CINES* et spécialement conçu pour l'archivage pérenne des données 3D en sciences humaines.

Fonctionnement d'un schéma de métadonnées structuré

Les métadonnées sont des informations que l'on écrit avec un formalisme très rigide.

Par exemple, il est nécessaire de connaître la langue qui a été utilisée dans l'archive : on décide alors d'utiliser la balise « *language* » de *DublinCore* (côté structure) qu'il faut renseigner avec la valeur

issue du code ISO-639-1 (côté valeurs), soit « fr » pour français. C'est ce couple (« *language* », « fr ») qui forme une métadonnée de l'archive. On peut définir ce couple (« *language* », « fr ») pour toute l'archive, ou pour certains fichiers seulement.

On pourrait également avoir besoin d'une information qu'aucun schéma connu ne propose :

« *ageDuCapitaine* » par exemple, qu'il faudrait renseigner par un nombre entier. Dans notre schéma nous pouvons alors créer cette nouvelle balise « *ageDuCapitaine* ». Cependant afin de maximiser l'interopérabilité (entre autres avantages), il est toujours prudent de réutiliser les schémas existants.

On renseigne les métadonnées d'un projet en se référant au schéma de structure. Le schéma de structure peut définir les niveaux et leur agencement :

- d'un côté, très concret : chaque fichier doit être décrit,
- de l'autre, plus abstrait : le projet d'ensemble doit être décrit,
- certains niveaux intermédiaires sont parfois utiles à décrire.

Pour ces niveaux, il s'agit de fournir 4 familles d'informations :

- **Le contenu utile** : ce qu'il y a dedans,
- **Le contexte** : comment, pourquoi ces données,
- **Les informations administratives** : droits, accès, propriété...
- **Les informations techniques** : taille du fichier, mode de création, validité.

Concrètement les métadonnées d'un projet peuvent s'écrire sous plusieurs formes. Nous faisons le choix de les rassembler dans un unique fichier. Ce fichier est une sorte de « dictionnaire » qui décrit l'archive et son contenu selon les modalités que nous définissons plus loin dans ce rapport. Souvent un logiciel avec une interface graphique assiste l'utilisateur dans la création de ce fichier, comme nous le décrirons également.

En conclusion, il s'avère que les services proposant d'archiver des modèles 3D (ou environnements virtuels) sont rares au regard des enjeux sous-jacents et inexistants en France. Pourtant de nombreux schémas de structure proposent des balises utiles. Cette problématique se retrouve au cœur de la politique de gestion des données à l'échelle européenne. Les travaux délivrés ici faciliteront la formulation des réponses à ce type de problématique.

2. INTRODUCTION À LA PRATIQUE DE L'ARCHIVAGE

Avant d'aborder l'intégration de l'archivage dans les processus de travail, la spécification des données à renseigner et la manière de les enregistrer, il est nécessaire de rappeler ce que la notion d'archivage recouvre exactement, et comment elle est mise en œuvre dans les domaines connexes à l'archéologie.

2.1. Définition

L'archivage désigne, selon la définition du *Larousse*, l'acte de collecte et de classement de données.

Ce terme est issu du bas latin *archivum* qui signifie « l'armoire pour les actes », mais il est lui-même dérivé du grec ἀρχεῖον *archeíon* qui signale un « bâtiment administratif, une magistrature » en relation avec les archontes²⁷. Le composant arche (ici préfixe comme dans architecture, mais parfois suffixe, comme dans monarchie, anarchie...) indique la notion de commandement ou d'autorité naturels, de commencement (par opposition au *cratos*, le pouvoir qui se conquiert, comme dans démocrate, bureaucrate).

L'archive n'est donc pas quelque chose de statique, mais elle est issue d'un processus (la collecte, le classement d'informations) et concerne un lieu en particulier (le meuble, le bâtiment administratif) en lien avec le pouvoir. Il s'agit donc des choses jugées importantes pour la vie de la société (l'ensemble des contrats établis entre différentes parties par exemple). De manière sous-jacente, la collecte induit un choix dans les informations à conserver et à la vue de la quantité de données, ce processus nécessite une gestion.

L'archivage numérique des représentations tridimensionnelles d'objets archéologiques n'échappe pas à ces principes généraux, et cela d'autant plus que la quantité de données produites dans ce contexte est souvent considérable.

2.2. Concepts liés à l'archivage numérique

De nos jours, la question de l'archivage est devenue floue en raison de la multiplication des documents produits, de la dématérialisation des supports et des nombreuses possibilités de stockage. En outre, le concept d'archivage a tendance à être confondu avec celui de la sauvegarde sécurisée. Ce serait une simple copie, à l'identique, de nos données sur un autre disque dur considéré comme plus fiable que celui utilisé pour travailler. En France, la réflexion sur l'archivage et sa mise en œuvre est confiée au *CINES* dès 2004 (*CINES* 2014a).

Comme le définit le *CINES* (*CINES* 2014b), « l'archivage recouvre aujourd'hui plusieurs sens, mais ce n'est pas une sauvegarde [...], ce n'est pas un service *HSM*²⁸ [...] et ce n'est pas l'ultime étape du stockage avant l'oubli ou la perte définitive » des données.

L'archivage d'un document numérique – tel que défini par le *CINES* – vise à le « conserver, le rendre accessible et en préserver l'intelligibilité » sur le très long terme, c'est-à-dire « plus de trente ans » (délai néanmoins extrêmement court au regard de l'âge des vestiges exhumés par l'archéologue). Le caractère informatique des données implique de détailler plus précisément les trois enjeux principaux de l'archivage :

²⁷ Les archontes sont des dirigeants politiques dans les cités grecques antiques.

²⁸ Hierarchical Storage Manager. Cela désigne un système permettant le stockage de données de façon hiérarchique. En fonction de différents critères (importance des données, rapidité d'accès nécessaire, etc.), les données sont réparties vers différents

supports (bande magnétique, disque dur, mémoire vive, etc.) Le système gérant l'intégralité de ce type de fonctionnement transfère les données en fonction des critères établis (liés à une contrainte de coûts du stockage par exemple). – Définition Wikipedia (site consulté le 24/07/2019).

- **Conserver le document dans le temps** : cela demande une gestion continue des supports physiques des contenus numériques puisque leur durée de vie est variable, et pour atteindre l'objectif des trente ans fixé par le *CINES*, plusieurs moyens de stockage successifs seront peut-être nécessaires. Satisfaire à cette nécessité échoit à l'organisme hébergeur de l'archive.

- **Assurer l'accès au document** : outre les droits de diffusion, il est important de toujours être en mesure de le lire. La donnée doit être enregistrée dans un format dont la structure est lisible et documentée, autrement dit un format ouvert, non-propriétaire. L'archivage consiste entre autres à valider les fichiers. Ils doivent être conformes au schéma auquel ils se réfèrent. Pour les modèles tridimensionnels, le travail conjoint d'*Archéovision* et du *CINES* a permis d'ajouter les formats *DAE* et *PLY* à la liste des formats de données pouvant être archivés.
- **Préserver l'intelligibilité du document** : être en mesure d'ouvrir et d'accéder au contenu d'un fichier n'est pas suffisant pour pouvoir le comprendre. Pour la compréhension et la confiance dans les données, il est indispensable d'accompagner l'archive d'informations permettant de comprendre dans quel contexte et comment celle-ci a été générée : ce sont les métadonnées. Par conséquent, les représentations tridimensionnelles en archéologie doivent être accompagnées de métadonnées décrivant l'objet concerné : par exemple, comment il a été numérisé, pourquoi, par qui, quels documents ont été mobilisés pour compléter la modélisation, etc.

Nous proposons ici, conjointement avec le *CINES*, un schéma de métadonnées permettant la préservation de l'intelligibilité de l'objet 3D

archivé et de donner les moyens à l'utilisateur d'utiliser ce schéma.

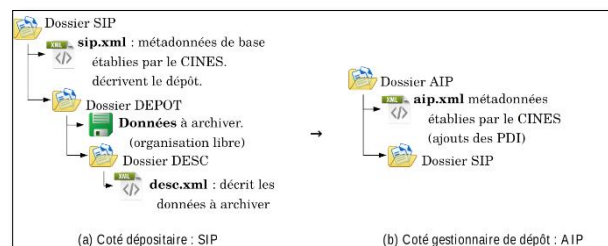


Figure 3 - Structure imposée par le CINES et création d'un AIP

2.3. Assurer l'intelligibilité des données

Le *CINES* dispose d'un schéma de métadonnées. Disponible en ligne, il sert à l'écriture d'un fichier au format XML qui accompagne obligatoirement les fichiers du dépôt dans le processus d'archivage.

Lorsque le besoin apparaît d'avoir une nouvelle métadonnée dans un schéma en construction, il est possible de récupérer une métadonnée provenant d'un schéma existant ou alors d'en créer une nouvelle ex-nihilo (voir partie 1.5). Partir d'une métadonnée existante facilite les relations entre schémas par la suite.

À ce jour, il existe un nombre croissant de schémas pour différents domaines. Un des plus anciens, et des plus utilisés, est le schéma *DublinCore* pour le monde du patrimoine. Le *CINES* utilise une partie de ce schéma dans le sien, celui que nous proposons en hérite également. Cette mutualisation de balises commune (« language » par exemple) facilite la maintenance et assure l'intelligibilité de la donnée associée, quelle que soit l'origine du lecteur.

Afin de conserver l'intelligibilité d'un document de type 3D, il est apparu nécessaire de mettre en place un schéma de métadonnées spécifiques à la 3D en SHS (métadonnées dites « métier »), qui complète le schéma du *CINES*. Ce nouveau schéma et son élaboration sont présentés dans les chapitres suivants.

2.4. Aspects techniques de la création d'une archive numérique

Une archive numérique rassemble plusieurs fichiers dans un seul dépôt. Afin d'assurer l'intelligibilité et la gestion des données, un ensemble d'information doit les accompagner sous la forme de métadonnées. Leur formalisation est définie par des packages.

Les packages

Le modèle *OAIS*²⁹ est un standard pour la gestion des archives numérique (voir figure ci-dessous). Le *CINES* l'utilise. Dans ce modèle, différents packages sont obligatoirement générés dans le cadre de la production d'archive : *SIP*³⁰, *AIP*³¹ et *DIP*³².

De notre côté, dépositaire de données à archiver, nous sommes concernés par le *SIP* que le *CINES* impose en fonction du type de soumission : il s'agit de décrire les données du dépôt et le dépôt en soi. Notre travail ici est une contribution à la création de métadonnées descriptives (*DESC*). Comme présenté en figure ci-dessus, ces métadonnées sont par la suite associées avec les métadonnées de bases du *SIP* imposées pour produire les données telles que stockées : le *AIP*. Le *AIP* est la combinaison de trois éléments principaux :

- Les éléments du *SIP* obligatoires du *CINES* qui renseignent la soumission du dépôt (en jaune sur la figure suivante). Nous incorporons ce point dans nos métadonnées.
- Les métadonnées « métiers » du *SIP* qui renseignent le contenu du dépôt (en bleu sur la figure ci-dessous). C'est l'objet du chapitre 4 et le cœur de nos travaux.
- Le contenu « Content Data Objects » qui constitue les données produites : le modèle 3D

et les sources (en jaune sur la figure ci-après). Ces éléments sont contraints par les formats définis en section 3.1. C'est la raison d'être de l'archive.

Le modèle *OAIS* sépare quatre composantes du « *Preservation Description Information* » (*PDI*) (figure ci-après) :

Reference : un identifiant unique et pérenne.

Provenance : l'historique de l'archive et description des contenus.

Context : relation entre objets, organisation des contenus de l'archive.

Access Rights Information : les droits associés à l'objet.

Fixity : démonstration de l'authenticité numérique (résultat d'un hachage)

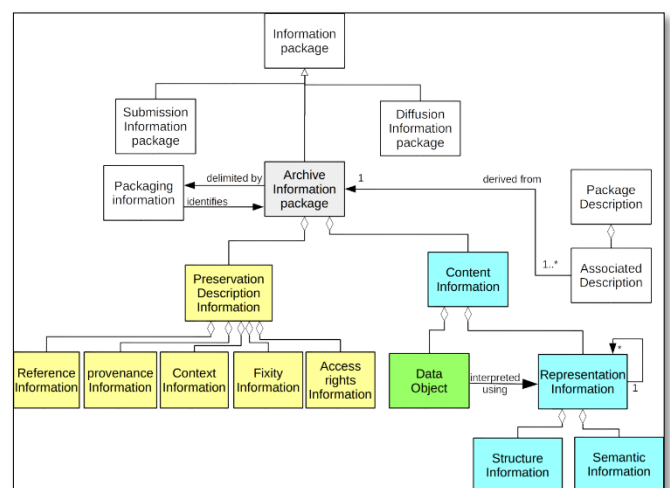


Figure 4 - Schéma UML centré sur le AIP, en jaune les parties imposées par le SIP de base du CINES ou complétées lors de la transformation en interne, en bleu les métadonnées sur lesquelles nous intervenons (*DESC*), en vert les contenus à archiver

²⁹ OAIS : Open Archival Information System

³⁰ SIP : Submission Information Package

³¹ AIP : Archival Information Package

³² DIP : Dissemination Information Package

La partie de droite (Package description) est donnée ici à titre indicatif. Elle concerne la diffusion des données, ce sont des vues produites pour indexer le *AIP* et faciliter sa recherche par un utilisateur (qui recevra alors le *DIP*). Ces vues sont produites en extrayant certaines informations du *AIP* en fonction des besoins du système de moissonnage (typiquement les métadonnées DublinCore). Nous ne sommes pas concernés par cette pratique. Nous notons que cette pratique d'associer la diffusion à l'archivage est plutôt anglo-saxonne.

Construction d'un SIP

La création du *SIP* incombe au dépositaire. Ce *SIP* doit satisfaire les exigences du gestionnaire d'archive.

« *European Archival Records and Knowledge Preservation* » (*E-ARK*) est un projet européen qui s'est terminé en 2017, visant à améliorer les méthodes et technologies d'archivage pérenne de données. Le patrimoine culturel et la 3D ne sont pas des préoccupations centrales. Néanmoins le projet s'est attaqué à la question de la production d'un *SIP* par un dépositaire non archiviste. Le développement d'un logiciel dédié à la création de packages d'archive (nommé *RODA-in*) marque un des aboutissements de ce programme. Ce logiciel permet d'associer à chaque fichier ou groupe de fichiers un ensemble de métadonnées qui doivent répondre à un schéma prédéfini.

RODA-in permet de décrire l'archive en conservant la structure et en ajoutant les métadonnées produites par l'utilisateur. Il est compatible avec les schémas classiques, et permet de créer des modèles de données hors schéma classique. Il permet de créer le *SIP* en *BagIt*³³ ou selon un formalisme propre au projet. *BagIt* est un système

très simple et largement utilisé qui conserve l'organisation hiérarchique des fichiers/dossiers et vérifie l'intégrité des contenus (empreinte de hachage). Ce système peut embarquer quelques descripteurs (métadonnées) pour l'ensemble du « bag ». De nombreux « bags » peuvent être créés, jusqu'à 1 par fichier.

```
bagExample/
|-- data
|   |-- projet34
|   |   |-- OCR
|   |       |-- source.png
|   |       |-- source.txt
|-- bag-info.txt
|   Bagging-Date: 2015-07-02
|   External-Identifier: urn:uuid:47556604-4a17-a39c-cdf752378feb
|   Contact-Email: matthieu@tinta.fr
|-- manifest-md5.txt
|   49afbd86a1ca9f34b677a3f09655eae9 data/projet34/OCR/source.png
|   408ad21d50cef31da4df6d9ed81b01a7 data/projet34/OCR/source.txt
|-- bagit.txt
|   BagIt-Version: 0.97
|   Tag-File-Character-Encoding: UTF-8
```

Figure 5 exemple de création du « bag »

³³ Pour plus d'information, cf. ([site](#) consulté le 25/07/2019)

3. ÉTABLIR UN PLAN DE GESTION DES DONNÉES POUR ARCHIVER

L'archivage des données produites au cours d'une recherche doit être anticipé. Il est impossible – ou du moins très fastidieux – d'atteindre cet objectif si les documents n'ont pas été préparés dans cette perspective. La raison est simple : renommer, convertir dans les formats acceptés, et nettoyer le contenu de plusieurs centaines de fichiers est une tâche chronophage et laborieuse.

Il conviendra donc de travailler dès le début du projet en suivant certaines règles. Cela consiste notamment à :

- Choisir des formats de fichier adaptés au projet et à l'archivage (obligatoire) ;
- Distinguer les fichiers qui seront archivés de ceux qui ne le seront pas (obligatoire) ;
- Renseigner les métadonnées (obligatoire) ;
- Adopter une arborescence de répertoires constante tout au long du projet (fortement conseillé) ;
- Nommer correctement les fichiers (fortement conseillé).

Tous ces points, détaillés ci-dessous, doivent être abordés puis résolus pour rédiger le plan de gestion des données (Cartier et al. 2015).

3.1. Formats de fichier

La mission du *CINES* est d'archiver les données reçues, avec comme perspective l'assurance de pouvoir les relire dans un avenir plus ou moins lointain. Or, les formats évoluent et l'interopérabilité des fichiers n'est pas garantie entre logiciels (versions). C'est pourquoi le *CINES* s'engage non seulement à assurer la pérennité du fichier sur un support physique, mais aussi à faire évoluer le contenu de celui-ci afin qu'il soit toujours conforme aux normes en vigueur. Cela

ne peut être possible que par l'utilisation de formats de fichiers documentés et les plus lisibles possible. Un nombre restreint d'extensions est ainsi disponible pour archiver au *CINES* les résultats de la recherche³⁴. Elles sont recensées ci-après en établissant une corrélation avec les types de données couramment manipulés en SHS.



Seules les extensions acceptées par le *CINES* dans une perspective d'archivage sont ici recensées ; cela n'interdit pas l'utilisation d'autres formats pour les phases intermédiaires de travail, mais ils ne pourront pas être intégrés à l'archive finale.

Pour les données issues de la numérisation

Cas Nature du fichier Format (extension)	Cas Nature du fichier Format (extension)	Cas Nature du fichier Format (extension)
Lasergrammétrie, topographie	Nuages de points	Préférer le Collada (* .dae), sinon le Stanford Polygon File Format (* .ply) en ASCII (et non pas en binaire).
Photogrammétrie	Clichés sans traitements	tiff, * .dng
Tomographie	Images en niveaux de gris associées à des descripteurs	Préférer le format ANALYZE (* .hdr et * .img), sinon DICOM

³⁴ Cf. ([site](#) consulté le 25/07/2019)

Pour les fichiers non-3D

Nature du fichier Format (extension)	Nature du fichier Format (extension)
Protocoles d'acquisition	*.odt, *.pdf, texte brut (*.tex,..)
Méthodologies (articles, mémoires...)	*.odt, *.pdf, texte brut (*.tex, *.txt...)
Archives utilisées	Audio : *.wav, *.oga (OGG), *.aiff, *.flac, *.aac Images : *.gif, *.jpeg, *.tiff, *.png, *.svg ; Texte : *.odt Vidéo : *.mp4 (MPEG-4), *.ogv (OGG), *.mkv
Autres textes (contrats, objectifs scientifiques...)	

Pour les fichiers 3D

Nature du fichier Format (extension)	Nature du fichier Format (extension)
Modélisation finale	*.dae, *.ply
Textures finales	*.gif, *.jpeg, *.tiff, *.png, *.svg

3.2 Organiser les données

La réalisation d'un dépôt d'archive consiste à identifier, puis rassembler, différents fichiers produits au cours de l'étude dans un même répertoire, puis à les renseigner. Aussi, il est important d'avoir une organisation cohérente des fichiers lorsqu'on élabore un modèle 3D afin de

faciliter, en fin de projet, la constitution de l'archive.

Cependant, en raison de la diversité des objets d'études et des méthodes de travail, il n'est pas possible d'édicter une règle universelle pour cette arborescence. En revanche, des exemples de pratiques tentant de concilier les contraintes du projet avec les perspectives d'archivage sont proposés en annexe 1.



[cf. annexe](#) – Cas d'étude : structurer son répertoire de travail

3.3 Nommer les fichiers

Un nom de fichier présente une information importante permettant – au premier coup d'œil – d'en identifier le contenu. Cela peut également être un moyen facilitant la structuration du répertoire de travail, surtout quand les fichiers 3D fonctionnent par écosystèmes (le fichier de géométrie avec les textures comme dépendance par exemple). C'est pourquoi cette question ne doit pas être éludée, mais anticipée dès la rédaction du plan de gestion des données.

Afin que les noms de fichiers soient compatibles avec la totalité des systèmes d'exploitation, certaines règles de construction peuvent être émises³⁵ :

- Seuls les caractères alpha-numériques doivent être utilisés, c'est-à-dire : les lettres `a → z`, les chiffres `0 → 9`, le tiret `` et le tiret bas `\_`. Aucune autre ponctuation ou caractère spécial ne peut être utilisé. Le tiret bas peut se substituer à une espace. Le point `` ne doit servir qu'à déterminer le début de l'extension du fichier.

- Un nom de fichier doit posséder une extension afin de faciliter l'identification de son contenu.

- Les caractères en majuscule et en minuscule ne sont pas toujours signifiants. Par

³⁵ cf., ([site](#) page 16 consulté le 25/07/2019)

exemple, les fichiers `rapport\_activite.odt` et `RAPPORT\_activite.odt` sont parfaitement identiques pour certains systèmes d'exploitation.

- Au sein d'une arborescence de répertoire correspondant à un projet, les noms de fichiers doivent être uniques afin d'éviter toute confusion au moment de la création des archives (tous les documents se retrouvent dans un seul dossier parent).

- La manière de nommer les fichiers devrait être persistante dans le temps et partagée par tous les acteurs du projet. Il est toujours intéressant d'expliquer les conventions de nommage des fichiers dans les parodonnées. Comme précédemment, la diversité des objets d'études, des méthodes de travail et des objectifs scientifiques et techniques, une règle universelle concrète ne peut pas être édictée. C'est pourquoi plusieurs cas d'étude réunis en annexe illustrent les pratiques.



[cf. annexe](#) – Cas d'étude : nommer les fichiers.

3.4 Sélectionner les données à archiver

Tous les documents produits au cours d'une étude n'intégreront pas un archivage pérenne. Pour des raisons de capacité de stockage et surtout de choix scientifiques, les fichiers intermédiaires ne sont pas à archiver. Il n'est pas non plus nécessaire de conserver ce qui pourra être reproduit ultérieurement, de manière plus performante (ex : le nuage de points produit en photogrammétrie).

En revanche, il est impératif de conserver : les données initiales, c'est-à-dire les sources, le résultat final, et les documents relatifs aux processus qui permettent de comprendre le passage de l'un à l'autre.

Plusieurs sources peuvent être utilisées pour créer et compléter un objet virtuel, que ce soit une numérisation 3D – lasergrammétrie, photogrammétrie ou tomographie – ou d'archives

– iconographies, textes, documents audiovisuels. Tous ces matériaux peuvent être combinés au sein d'un même système de représentation tridimensionnelle pour obtenir un modèle cohérent. Les informations qui auront servi à représenter directement ou indirectement un état de l'objet virtuel doivent impérativement être présentes dans le dépôt. Le document qui les présente peut lui-même être numérisé ou inséré sous la forme d'un identifiant pérenne (un numéro d'identification dans une base de données accessible). Il est important de conserver ce lien entre les sources utilisées pour la numérisation et la modélisation finale, car c'est ce qui constitue une partie de l'argumentation, et confère le rang de démarche scientifique à l'ensemble du processus.

Le modèle 3D final est au cœur du dépôt. Tous les fichiers présents dans le dépôt servent à comprendre ce document. Selon la nomenclature établie au sein du Consortium 3D SHS, un modèle 3D peut se voir attribuer un numéro de version (V0, V1 ou V2 – cf. le tableau des pratiques 3D en SHS). Il faut souligner que c'est bien la version finale qui doit être archivée et non pas la version n-1, car elle présente une validation moins importante ; elle possède moins de valeur ajoutée que le dernier état.

La transformation des sources en modèle 3D à archiver nécessite des étapes plus ou moins nombreuses, plus ou moins complexes. Ces fichiers intermédiaires, stockés sur un disque dur local, ne participeront pas à l'archive. En revanche, les processus peuvent être décrits, et il est important que cette documentation soit accessible sous la forme de parodonnées : texte libre (formats odt, ...), tableaux (formats csv, ...). Deux catégories sont identifiées : les processus techniques et les choix de restitution.

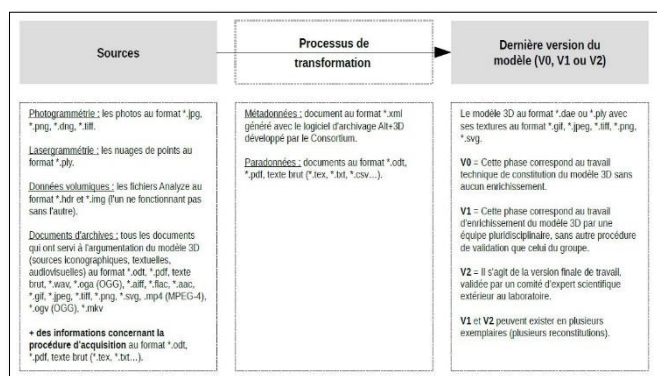


Figure 6 Schéma identifiant les données à archiver

D'une manière générale, les documents qui ne pourront pas être réalisés à nouveau sont archivés :

- **Pour un modèle de niveau V0** : les données primaires (nuages de points de lasergrammétrie - mais pas de photogrammétrie - clichés, etc.) et les données nécessaires pour refaire une modélisation du terrain (texte expliquant les objectifs scientifiques et techniques recherchés, aspects techniques, etc.).
- **Pour un modèle de niveau V1 et plus** : les données initiales des différentes sources utilisées, le cheminement intellectuel conduisant à l'hypothèse de restitution, la restitution elle-même.

3.5 Renseigner les données

Les informations peuvent être renseignées sous deux formes différentes : sous la forme de métadonnées et de paradonnées. Si les métadonnées sont conçues pour être de forme courte et comprise dans un schéma structuré, les paradonnées sont beaucoup plus libres afin de mieux s'adapter aux contraintes du projet.

Paradonnées

Les paradonnées, renseignent le « comment » des données du dépôt. Il s'agit de fichiers à part entière, traitant par exemple des processus techniques et des choix de restitution :

- **Les processus techniques** : les solutions techniques choisies peuvent avoir un impact non négligeable sur la qualité du modèle 3D produit. Par exemple, les algorithmes permettant de créer un maillage peuvent produire des résultats de qualité variable et donc avoir un impact important sur les données produites. Il pourrait donc être intéressant de restituer l'ensemble de ces choix technologiques afin qu'une analyse critique puisse être formulée. Au cours du traitement des données, il est possible d'utiliser des logiciels qui proposent de générer automatiquement des rapports. Par exemple, *Agisoft Photoscan* propose de produire une documentation du projet en PDF et en XML (données stockées dans l'arborescence des dossiers). Il est fortement conseillé d'inclure ce type de documentation au sein de l'archive.
- **Les choix de restitution** : la restitution d'un objet archéologique est toujours le résultat d'une pensée. Son élaboration peut être plus ou moins complexe, et est le résultat de choix successifs. Il s'agit donc de rendre compte des arguments qui ont servi à privilégier telle solution plutôt que telle autre. Cela se matérialise dans sa forme la plus développée par un texte argumenté. Moins formalisés, des comptes rendus de réunion restituant la nature des échanges – voire les courriers électroniques – peuvent être utiles, mais il faudra néanmoins veiller à ce que la réflexion soit suffisamment explicite.

La diversité des objets d'étude, des objectifs scientifiques et des techniques mobilisées exclut de présenter une fiche-type de renseignement. C'est pourquoi ce point sera traité sous la forme d'exemples en annexe.



cf. annexe – Cas d'étude : rédiger les paradonnées du projet

Métadonnées

Les métadonnées, renseignent le « quoi », le contenu des données du dépôt. Elles sont généralement courtes et factuelles (exemples : titre, auteur, date) Cf partie 1.4.

Le fichier contenant la 3D, ainsi que l'écosystème de fichiers mobilisés pour sa constitution et son argumentation, présentent de nombreuses singularités dues au format 3D des données produites, qui peuvent être documentées sous la forme de métadonnées. Afin de répondre de manière la plus pertinente possible à cette problématique, un schéma spécifique a été rédigé. Il est détaillé dans le chapitre suivant.

3.6 Conclusion

Pour rendre possible l'archivage des productions scientifiques 3D, un plan de gestion des données doit être rédigé. Cela permet de produire des informations directement compatibles avec les impératifs de l'archivage. Tous les aspects d'un PGD n'ont pas été abordés ici, seulement ceux qui méritent une attention particulière pour répondre aux contraintes rencontrées avec les modèles 3D et leur archivage par le *CINES*.

4. MÉTADONNÉES NÉCESSAIRES À L'ARCHIVAGE DES MODÈLES 3D

4.1 Méthodologie mise en œuvre pour définir ces métadonnées

Le contexte (cf. le chapitre 1.3) indique que la réflexion sur les métadonnées permettant de documenter les objets archéologiques n'est pas nouvelle. Un schéma rédigé par le consortium britannique *ADS – MIDAS* – a été publié dès 1998. Il s'est étoffé avec le temps de manière à couvrir un grand nombre d'aspects de la documentation des objets patrimoniaux. Nous assumons le choix de ne pas nous être appuyés sur ce standard pour définir un nouveau modèle de données. Notre volonté était de produire le schéma le plus léger et le plus succinct possible pour ne pas alourdir la procédure de constitution du dépôt d'archive. Cela s'est traduit par la non-considération des schémas existants. Néanmoins, le schéma correspond aux orientations définies par le *CINES*. Il s'adosse ainsi à un haut niveau d'exigence concernant l'archivage et répond sans superflu aux exigences métiers en archéologie.

Le premier temps de notre travail a consisté à comprendre le schéma *PAC* (Plateforme d'Archivage au *CINES*) et à le « configurer » pour qu'il devienne le cadre général de réflexion. Son objectif est de permettre la compatibilité du dépôt d'archive avec les exigences du *CINES*. En somme, cette partie intéressera davantage le gestionnaire de la documentation que l'archéologue. Cette structure n'est pas définitive et sera bien entendu amenée à évoluer. Si nous nous sommes appuyés sur la dernière version (numéro 2.3)³⁶, il faudra veiller à assurer la conformité de cette partie avec le standard défini par le *CINES*.

Une deuxième catégorie d'informations a ensuite été définie : celle permettant de décrire le ou les objets physiques concernés par le dépôt. Cet ensemble concerne l'ensemble de l'archive puisqu'il a été tranché que le contenu correspondait à un groupe cohérent d'entités – dont la composition est laissée à la libre appréciation de l'archéologue. Ainsi, ce sont des données qui intéressent directement l'archéologue.

Enfin, un dernier ensemble de métadonnées a été défini : celui permettant de décrire les représentations tridimensionnelles. La réflexion s'est dans un premier temps appuyée sur un document interne à Archéovision qui listait des informations jugées importantes pour documenter les objets 3D. Cette ébauche a été amendée, allégée et restructurée de manière à être cohérente avec les autres parties du schéma. Un soin important a aussi été accordé à la manière de renseigner les entités archéologiques et de les associer avec les sources ayant permis de les modéliser, les documenter, et les résultats scientifiques obtenus. L'articulation entre ces données de différentes natures et les deux ensembles précédents a été résolue en créant le concept d'objet virtuel. Il ne concerne pas un fichier 3D, mais regroupe un ou plusieurs maillages présents dans un ou plusieurs fichiers – la pertinence de l'association étant définie par le créateur de l'archive. Ce sont, cette fois, des métadonnées plus techniques qui devraient être prioritairement complétées par celui qui a réalisé les modèles 3D.

³⁶ cf. ([site](#) consulté le 24/07/2019)

Ces différents objectifs ont été poursuivis et accordés de manière à produire une liste de métadonnées nécessaires à l'archivage des modèles 3D en archéologie. À l'occasion de cette synthèse, plusieurs niveaux d'utilisateurs ont été définis : le gestionnaire de la documentation, l'archéologue et l'infographiste.

4.2 Schéma proposé

Le schéma (figure ci-dessous) est construit autour de trois pôles : le projet 3D (*zone grise*), l'objet virtuel (*zone rouge*) et les données le renseignant (*autres couleurs*) – avec une distinction entre l'écosystème de fichiers 3D (*zone bleue*) et les documents sources (*zone verte*). Chacun de ces pôles est composé d'une ou plusieurs classes.

L'objet virtuel (**objetVirtual**) est le cœur du schéma puisque c'est autour de cette notion que tout s'agence. Le projet 3D constitue une introduction afin de définir le site, les objectifs de l'étude menée et les caractéristiques du dépôt (**depot:Archeo**, **depot:gestion** et **depot:structure**). Les documents sources permettent de constituer et de renseigner l'objet virtuel en distinguant la 3D (**maillage**, **fichier3DGeometrie** et **fichier3DTexture**) des sources (**fichierArchive**, **fichierDonneeVolumique**, **fichierLasergrammetrie**, **fichierParadonnee** ou encore **fichierPhotogrammetrie**).

Chacune de ces classes regroupe un certain nombre de descripteurs. Ces derniers sont décrits dans le catalogue de métadonnées (cf. Annexe). Les relations entre les différents groupes sont explicitées à travers une modélisation UML (cf. l'illustration X). Mais définissons maintenant les différentes catégories :

depot:Archeo permettent de renseigner diverses informations au niveau du site archéologique concerné par le dépôt et le cadre dans lequel le projet 3D a été conduit.

depot:gestion regroupe un ensemble de métadonnées permettant de comprendre le statut

du dépôt : qui l'a réalisé, les droits associés au contenu, etc.

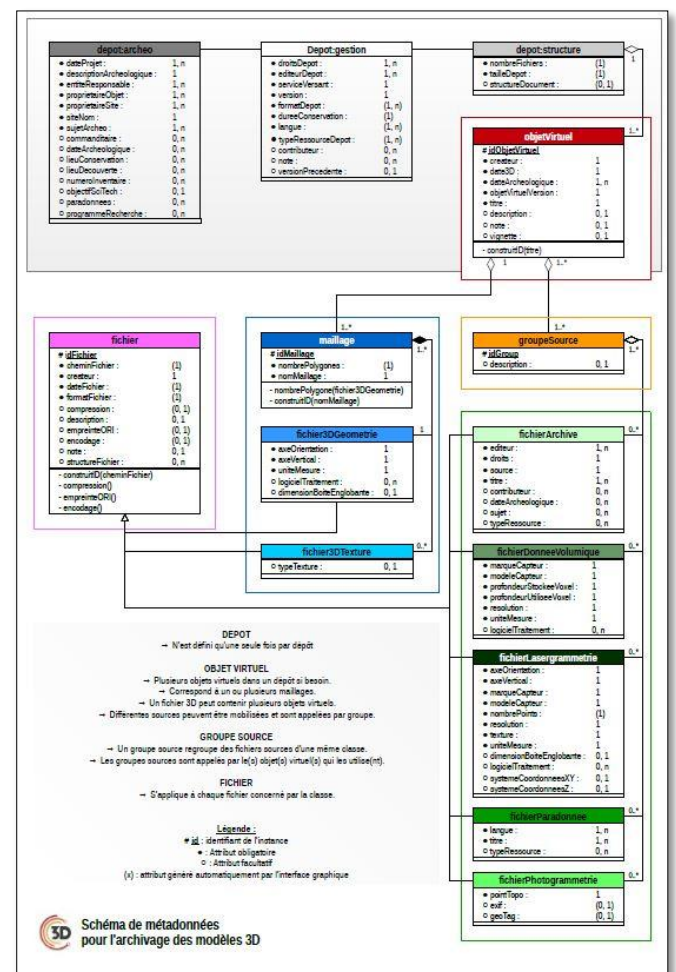


Figure 7 Schéma de métadonnées pour l'archivage des modèles 3D

depot:structure a pour fonction de clarifier la manière dont les fichiers présents dans le dépôt s'agrent et fonctionnent entre eux.

objetVirtual est un objet archéologique (notion volontairement floue). Il est composé d'un ou plusieurs maillages (**maillage**), et il est documenté par un ensemble de sources, appelé **groupeSource**.

Un **maillage** ne peut être contenu que dans un seul fichier de géométrie (**fichier3DGeometrie**). Plusieurs textures (**fichier3DTexture**) peuvent s'appliquer sur lui. C'est la seule classe mobilisée pour la description 3D directe de l'objet virtuel.

Les fichiers de géométrie (**fichier3DGeometrie**) peuvent contenir un ou

plusieurs maillages (``maillage``). Note : le *.ply n'accepte qu'un seul maillage par fichier, contrairement au *.dae.

Plusieurs fichiers textures (``fichier3DTexture``) peuvent s'appliquer sur un même maillage (``maillage``). Ils sont alors regroupés par tag.

``groupeSource`` est une classe qui documente l'``objetVirtuel``. Il permet de regrouper les différentes sources possibles : ``fichierArchive``, ``fichierDonneeVolumique``, ``fichierLasergrammétrie``,

``fichierParadonnee`` ou encore ``fichierPhotogrammetrie``.

``fichierArchive`` concerne des fichiers de n'importe quel type, qu'il s'agisse d'images, de textes, ou de témoignages audiovisuels. Ce sont des données d'archive, des sources historiques qui renseignent un état antérieur de l'objet virtuel et qui n'existe plus. En ce sens, les documents sources sont indispensables dans le cas des restitutions. Ce sont des fichiers qui peuvent porter l'extension *.wav, *.oga (OGG), *.aiff, *.flac, ou *.aac pour l'audio ; *.gif, *.jpeg, *.tiff, *.png, ou *.svg pour les images ; *.mp4 (MPEG-4), *.ogv (OGG), *.mkv pour les vidéos ; *.odt pour les textes s'ils ont été numérisés. Cette classe alimente ``groupeSource``.

``fichierDonneeVolumique`` concerne les acquisitions qui produisent des voxels. Deux fichiers sont nécessaires pour ce type de données : l'un à l'extension *.img, l'autre étant *.hdr. Un tag permet de grouper ces deux fichiers décrivant la même acquisition à l'objet virtuel. Cette classe alimente ``groupeSource``.

``fichierLasergrammétrie`` regroupe également des fichiers du type nuage de points obtenus par lasergrammétrie, exclusivement au format *.ply. Cette classe alimente ``groupeSource``.

Les paradonnées (``fichierParadonnee``) sont des fichiers de n'importe quel type dont le contenu permet de mieux comprendre le processus de création de l'objet virtuel. Les informations peuvent renseigner à la fois l'acquisition, le traitement, et l'argumentation archéologique dans le cas des restitutions. Sans pouvoir être exhaustif, il est possible d'inclure des documents au format *.odt³⁷, *.pdf³⁸ ou encore *.xml³⁹. Cette classe alimente ``groupeSource``.

Enfin, ``fichier`` regroupe un ensemble de métadonnées qui s'applique à tous les fichiers contenus dans le dépôt. C'est pourquoi elle vient compléter les classes ``fichier3DGeometrie`` et ``fichier3DTexture``) ``fichierArchive``, ``fichierDonneeVolumique``, ``fichierLasergrammétrie``, ``fichierParadonnee`` et ``fichierPhotogrammetrie``.



cf. [annexe](#) – catalogue des métadonnées.

4.3 Philosophie du modèle de données

Le modèle de données proposé :

- **Est le plus à plat possible** : les groupes des métadonnées sont peu imbriqués.
- **Est le moins contraint possible** : l'utilisateur est libre de renseigner l'information sous la forme de son choix : une chaîne de caractères. Le dépositaire est responsable de l'intelligibilité, des conseils sont proposés.
- **Est le plus léger possible** : la plupart des descripteurs sont facultatifs.
- **Offre une gestion souple du projet** : le(s) fichier(s) 3D contenant plusieurs entités archéologiques / maillages / objets virtuels.
- **Est aligné sur les standards dès que possible**. Cela permet d'être moins ambigu puisque les descriptions sont claires dans le cas des modèles 3D.

³⁷ Des notes prises pendant la confection des modèles 3D, la rédaction des résultats scientifiques de l'étude, etc.

³⁸ Le rapport généré automatiquement par Photoscan par exemple.

³⁹ Les fichiers produits par Photoscan dans son répertoire de travail et qui renseigne la position, l'orientation, et la calibration de chaque caméra dans son répertoire de travail.

- Permet une organisation indépendante de l'organisation fichier/dossier de l'utilisateur. Les objets virtuels sont au cœur du dépôt. Ce sont des objets archéologiques et appellent :

- Un fichier DAE (qui peut être appelé par plusieurs objets virtuels) ;
- Un groupe de fichier de texture (via un tag) ;
- Un ou plusieurs groupes de fichiers sources via des tags. Par exemple deux ensembles de photos (tag mesphotos1 et cesphoto2), 3 ensembles documents d'archives (tags archives1, archs2 et trucs3) et un LIDAR (tag monlidar).

4.4 Perspectives

Le schéma tel qu'exposé a servi de base pour rédiger le document maître XSD ; celui-ci est indispensable pour générer les fichiers *.xml qui définiront le dépôt d'archive. Le plus grand soin a été apporté pour que le produit soit compatible avec les standards du *CINES* – et ce même si un autre service que le *CINES* était retenu pour archiver les données.



cf. mdacst3D.xsd⁴⁰

Toutes les techniques d'acquisition 3D ne sont pas prises en compte dans ce schéma⁴¹, qui ne comprend pas non plus la diversité des fichiers pouvant être utilisés pendant une étude archéologique⁴² – et encore moins par conséquent dans le domaine des sciences humaines et sociales. Les efforts se sont uniquement portés sur les modèles 3D et l'argumentation scientifique à partir de ceux-ci. La limitation volontaire des procédures d'acquisition se justifie par le manque de recul dans la pratique des auteurs du schéma pour être certain d'être pertinent dans le choix des métadonnées.

Ce schéma possède néanmoins un potentiel d'évolution puisque sa structure ne présente aucune imbrication. Il sera facile d'ajouter une nouvelle classe alimentant le `tagSource` pour inclure de nouvelles techniques d'acquisition par exemple, et en attribuant des descripteurs existants – et des nouveaux – à ce module.

Ces changements sont en cohérence avec la politique du *CINES* qui autorise les évolutions dans les schémas de métadonnées métier.

⁴⁰ <http://altag3d.huma-num.fr/data/mdacst3d.xsd>

⁴¹ Exemple : les images hyperspectrales, les acquisitions en lumière structurée, la thermographie, etc.

⁴² Exemple : les bases de données, les données SIG, etc.

5. CONSTITUTION DE L'ARCHIVE

5.1 Procédure pour constituer les métadonnées

Les métadonnées sont des informations renseignant les données de l'archive, enregistrées dans un fichier au format **.xml*. Or, il est indispensable d'utiliser une interface utilisateur pour écrire ce type de fichier. La réflexion a donc été poussée jusqu'à la manière de remplir le plus simplement possible toutes les informations décrites dans le chapitre précédent, tout en visant la création automatique du dépôt d'archive compatible avec les spécifications du *CINES*.

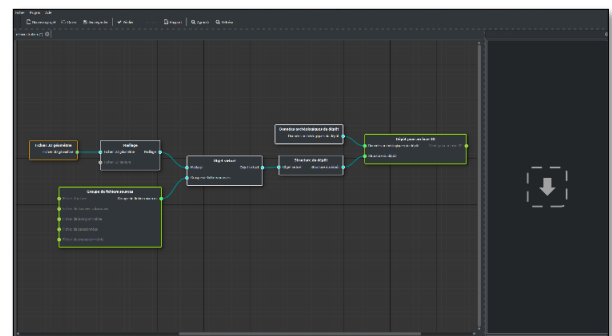
Lorsque nous avons voulu confronter le schéma de métadonnées à un exemple concret, il est apparu qu'un grand nombre de relations entre les différents types de documents sources, les fichiers 3D, et les objets virtuels ne pouvaient pas être renseignés avec les interfaces classiques ressemblant à des formulaires. Une toute autre stratégie a alors été adoptée. Le principe s'appuie sur une représentation visuelle des paquets de fichiers à archiver ; ces paquets sont reliés entre eux pour retrouver la hiérarchie du dépôt. Ainsi, la structure du schéma *UML* qui a été présentée a été utilisée pour ainsi dire in extenso pour la rédaction du schéma *XSD* et la formalisation du logiciel.

Le résultat de cette réflexion est le logiciel *aLTAG3D*. C'est une interface mettant en avant les relations entre les différentes sources de données. La personne souhaitant archiver ses données 3D devra, par un simple glisser-déposer, déposer les fichiers qu'il souhaite archiver dans l'espace de travail de l'application et établir les relations entre les différents fichiers. Certaines informations seront extraites automatiquement des fichiers déclarés dans le projet d'archivage. Le reste devra être rempli par l'utilisateur. Le logiciel se charge

ensuite de générer les fichiers **.xml* des métadonnées.

Le logiciel *aLTAG3D*⁴³ a été défini avec les contraintes suivantes :

- Un logiciel multi-plateforme, capable de générer le dépôt dans son intégralité, en plusieurs fois si besoin (enregistrement du travail en cours, export éventuel sur un autre ordinateur).
- Un design mettant en avant les liens entre les informations à compléter et en soulignant celles qui sont à renseigner obligatoirement.
- Une interface basée sur le glisser-déposer.
- Une structure intégrant le schéma de données métier, bâti par le Consortium 3D et capable d'intégrer automatiquement les mises à jour.
- Une documentation accessible depuis l'interface du logiciel.
- Une licence open source.
- Un rapport de fin accompagnant la constitution de l'archive.



cf. Site officiel du logiciel⁴⁴



cf. la documentation "utilisateur" du logiciel⁴⁵



cf. la documentation "développeur" du logiciel⁴⁶

⁴³ aLTAG3D: a Long Term Archive Generator for 3D models.

⁴⁴ Cf. ([site](#) consulté le 24/07/2019)

⁴⁵ Cf. ([site](#) consulté le 24/07/2019)

⁴⁶ Cf. ([site](#) consulté le 24/07/2019)

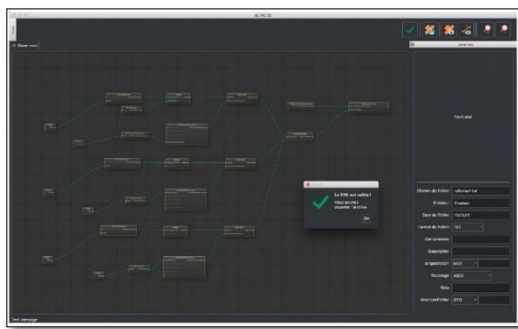
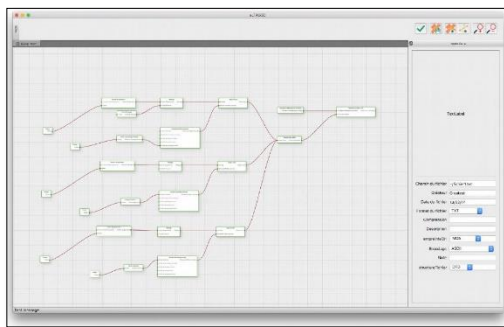


Figure 9 aLTAG3D (thème sombre)



5.3 Procédure pour constituer les parodonnées

Les parodonnées sont des documents texte au format *.odt ou *.pdf. La création de ce type de fichier est donc déjà maîtrisée pour des chercheurs en SHS. En raison de la diversité des objectifs scientifiques et techniques poursuivis, il n'est pas possible d'établir une liste de recommandations précises. Le choix du type d'information à renseigner et la manière de les structurer sont laissés à la discrétion du chercheur. Il sera donc de la responsabilité du déposant de veiller à ce que toutes les informations utiles à la réappropriation des documents soient disponibles.

Néanmoins, des récurrences peuvent apparaître dans la constitution des parodonnées, tant du point de vue de la structure que de la nature des informations indispensables à la bonne compréhension des données. Une enquête auprès de la communauté francophone en SHS a été réalisée en juin 2017 afin de recenser les différentes pratiques sur ce sujet. Les réponses les plus pertinentes sont collectées en annexe.

[cf. annexe](#) – Rédiger les parodonnées du projet

5.2 Finaliser un dépôt

Une fois que l'ensemble des fichiers est renseigné dans *aLTAG3D*, et que le projet est valide, un rapport peut être généré, mais surtout il est possible de créer l'archive dans le répertoire de son choix. Cela consiste à copier toutes les données et à écrire tous les fichiers. *XML selon une structure compatible avec les attentes du CINES. Il faut encore une fois souligner que le CINES n'est pas le seul destinataire possible pour archiver les données, mais que n'importe quelle structure peut s'engager au début d'un programme de recherche à conserver les données produites selon des modalités qui doivent être définies par les deux parties. L'utilisation du logiciel *aLTAG3D* garantit seulement que les données sont structurées selon les normes du CINES, reconnu pour son haut niveau d'exigence dans ce domaine.

Une fois l'export finalisé, le répertoire peut être encapsulé dans un fichier archive dont l'extension peut-être *.zip, *.7z, *.tar, *.rar, etc. Cela facilitera l'envoi du fichier vers le service en charge de l'archivage.

6. CONCLUSION ET PERSPECTIVES

Dans un premier temps, les enjeux et les spécificités propres à l'archivage ont été présentés. À partir des spécifications du *CINES* – structure de référence pour l'archivage des données de la recherche en France – et des travaux initiés par *Archéovision* (Bordeaux), un schéma de métadonnées a été construit, et chaque clé a été détaillée. Cependant, cette manière de documenter ne pouvait pas intégrer toutes les informations nécessaires à une future exploitation ; l'importance des paradosées a alors été soulignée.

L'objectif à terme est de publier ces recommandations sous la forme d'un nouveau schéma spécifiquement dédié à l'archivage pérenne des modèles 3D. Il serait alors souhaitable que ce document soit traduit en diverses langues pour qu'il puisse profiter à la communauté européenne – voire internationale. Cet objectif sera également atteint lorsque le schéma produit par le *Consortium 3D SHS* pour l'archivage des modèles 3D sera mappé sur d'autres schémas existants (*MIDAS*, *METS* par exemple), pour assurer une meilleure définition des métadonnées à renseigner. En s'appuyant sur ces standards, la marge d'interprétation de l'information attendue sera moins grande, et le document sera plus facilement adopté pour de futures évolutions.

Néanmoins, il est déjà reconnu qu'il n'est pas exhaustif et qu'il devra évoluer pour intégrer des techniques d'acquisition sur lesquelles nous n'avons pas encore d'expérience. Nous pouvons citer par exemple les acquisitions en lumière structurée, les images hyperspectrales, thermiques, etc. De plus, il ne répond qu'aux attentes propres à l'archéologie pour le moment,

mais le schéma est suffisamment souple pour intégrer de nouvelles pratiques. En l'occurrence, les projets de modélisation et d'informations géographiques sont en train de converger vers les mêmes outils de travail, ce qui provoque l'émergence de nouveaux standards. En cours de constitution et portés par *Open Geospatial Consortium*⁴⁷, ils pourront à terme être intégrés dans le schéma que nous portons – nonobstant la création d'un validateur auprès du *CINES*.

Il faut aussi souligner que, s'il a été décidé de fonder les pratiques d'archivage sur les recommandations exigeantes du *CINES*, la volonté a été de produire un schéma léger, souple d'utilisation et facile à faire évoluer, adoptable par n'importe quelle structure. Si un autre partenaire pour l'hébergement est choisi, la solution restera néanmoins compatible avec les standards actuels.

De cette manière, la méthode proposée par le *Consortium 3D SHS* peut être déployée par n'importe qui – un chercheur comme un acteur privé – et à différentes échelles – au sein du laboratoire, en partenariat avec une administration, ou sur les serveurs du *CINES*⁴⁸.

Au regard de cette possibilité à stocker de manière pérenne les dépôts d'archive, le logiciel *aLTAG3D* a été développé. C'est une interface utilisateur qui analyse un document maître (le fichier *XSD* du schéma de métadonnées). Ainsi, l'arborescence qui pourra être créée dépend entièrement du contenu de ce document maître. Le logiciel est conçu de manière à être très flexible, ce qui laisse envisager une déclinaison de son utilisation pour intégrer d'autres documents maîtres : il suffira d'indiquer par exemple la nature des données qu'on souhaite archiver pour mobiliser la structure adéquate et générer le dépôt de la même manière.

⁴⁷ Cf. ([site](#) consulté le 24/07/2019).

⁴⁸ Rappelons que le *CINES* est la seule structure qui s'engage à faire évoluer le contenu des fichiers selon l'évolution des formats de fichier afin d'assurer leur lisibilité dans le temps. Une

administration peut avoir les capacités techniques pour assurer la pérennité des fichiers dans le temps. Ces objectifs sont en revanche plus difficiles à atteindre au sein d'un laboratoire ou d'une entreprise.

7. CE QU'IL FAUT RETENIR...

... DU POINT DE VUE DE CELUI QUI VEUT ARCHIVER SES DONNÉES

1. L'archivage doit être pensé dès le début du projet, car de nombreux points doivent être anticipés. La façon la plus optimale de procéder est de rédiger un plan de gestion des données.
2. L'archivage devrait être réalisé tout au long de l'étude, et non à la fin.
3. Le répertoire de travail devrait être stabilisé dès le début du projet. Cela permet de constituer le projet d'archive au fur et à mesure (c'est le chemin du fichier qui est enregistré, et non le fichier lui-même ; en le déplaçant, le lien est rompu).
4. Deux fichiers ne peuvent pas partager le même nom dans le dépôt d'archive. Par conséquent, il faut réfléchir à une manière de nommer les documents qui ne génère pas ce type de conflit. Les noms ne doivent inclure ni caractères spéciaux ni espaces.
5. Tous les formats de fichier ne peuvent pas être archivés. Par conséquent, il faut se rapprocher des préconisations du *CINES* pour assurer une intelligibilité des données à long terme.
6. Tous les fichiers produits au cours d'une étude ne seront pas archivés, seulement une sélection d'entre eux :
 - les données issues de la numérisation (aux formats compatibles avec le *CINES*),
 - la procédure décrivant la transformation de ces données (paradonnées),
 - les archives associées à l'objet d'étude,
 - le modèle 3D final de l'étude.

Afin de guider le futur utilisateur dans l'archivage des modèles 3D, des exemples illustrent en annexe la résolution de certains points par d'autres équipes de recherche.

... AU SUJET D'ALTAG3D

1. *aLTAG3D* est une interface utilisateur basée sur la librairie *Qt*. Cet environnement permet de ne coder qu'une seule fois le logiciel et de le déployer vers les trois plateformes de travail le plus courantes (Windows, Mac OSX et GNU Linux).
2. Le logiciel est diffusé sous une licence libre (*GPL*).
3. Cette interface établit un lien entre l'utilisateur et le document maître *XSD*. Ce dernier n'est pas inclus dans le logiciel ; il est consulté via une interrogation sur le serveur où il est stocké.
4. L'arborescence des données créée et visualisée à travers l'interface est directement issue de la structure du document *XSD*. Par conséquent, l'utilisateur accédera toujours à la dernière version du schéma, il ne sera pas nécessaire d'effectuer une mise à jour logicielle.
5. En d'autres termes, le logiciel développé pourrait accepter d'autres schémas de données avec relativement peu d'adaptation.
6. Le logiciel a été conçu de manière à visualiser la progression de la création du dépôt grâce à un code couleur. Un système de validation permet de vérifier la validité du projet avant d'accéder aux fonctions d'export.
7. L'ajout des fichiers dans le projet d'archive ne déplace pas les fichiers concernés. Ce n'est qu'à la génération de l'archive que les fichiers sont copiés dans le répertoire indiqué.
8. *ALTAG3D* est en mesure de générer un rapport regroupant les principales caractéristiques de l'archive au format *HTML*.

8. BIBLIOGRAPHIE

ADS, 2014. *Guidelines for Depositors*, version 2.0, September 2014 ([site](#) consulté le 24/07/2019)

Cartier A., Moysan M., Reymonet N., 2015. *Réaliser un plan de gestion de données de la recherche*

Data management Plan (DMP), guide de rédaction (V1, 09/01/2015). Cf. ([site](#) consulté le 24/07/2019)

CINES 2014a. *Le concept d'archivage numérique pérenne*. CF. ([site](#) consulté le 24/07/2019), dernière mise à jour le 22/04/2014

CINES 2014b. *Une plateforme en production* ([site](#) consulté le 24/07/2019), dernière mise à jour le 23/06/2014

Denard, Hugh "A New Introduction to the London Charter" in A. Bentkowska-Kafel,

D. Baker & H. Denard (eds.) *Paradata and Transparency in Virtual Heritage Digital Research in the Arts and Humanities Series* (Ashgate, 2012) 57-71. <http://www.londoncharter.org/> (site consulté le 26/06/2019)

Perrin K., Brown D. H., Lange G., Bibby D., Carlsson A., Degraeve A., Kuna M., Ylva Larsson Y., Pálsdóttir S. U., Stoll-Tucker B., Dunning C., Rogalla von Bieberstein A., 2014. *Référentiel et guide des bonnes pratiques pour l'archivage archéologique en Europe – EAC Guidelines 1*. Europae Archaeologiae Consilium, Namur, Belgique, 70 p. ([site](#) consulté le 26/07/2019)

Rocheleau M., 2010. *La modélisation 3D comme méthode de recherche en sciences historiques*. Actes du 10ème colloque International étudiant du département d'histoire, Université Laval, Québec, pp. 246-265. www.academia.edu ([site](#) consulté le 26/07/2019)

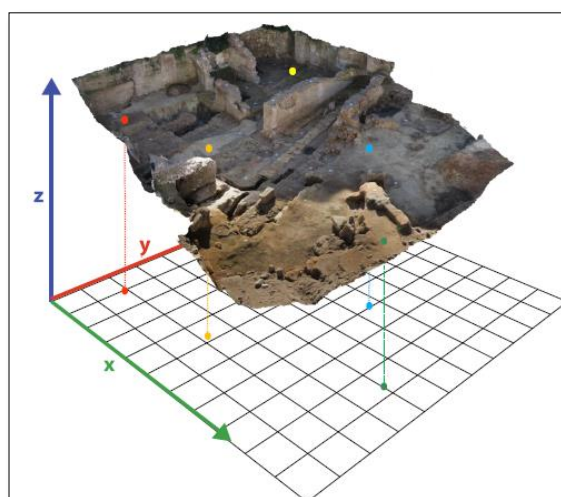
The Spanish Society of Virtual Archaeology (SEAV) 2011. *Principles of Seville*, International Forum of Virtual Archaeology, ([site](#) consulté le 25/07/2019)

RECOMMANDATIONS DU CONSORTIUM 3D SHS

SYNTHÈSE DES OUTILS ET DES TECHNOLOGIES 3D

APPORT CRITIQUE SUR LES MATÉRIELS ET LOGICIELS 3D

ÉDITION 2019



VERSION 2019

Anthony Pamart (CNRS, MAP - UMR 3495)

VERSION 2018

Violette Abergel (CNRS, MAP - UMR 3495)

Anne Flammin (MOM FR3747, CNRS, ArAr UMR 5138)

Charlie Morineau (UFR HSS - CIREVE)

Hervé Paitier (INRAP)

Anne Schmitt (MOM FR3747, CNRS ArAr 5138)

Sabine Sorin (CNRS, CEPAM UMR 7264)

1. AVANT PROPOS

1.1 Complémentarité technique

Les flux de production de modèles 3D, qu'ils soient issus d'une modélisation ou d'une numérisation sont généralement résultants d'un processus complexe pouvant s'appuyer sur différents supports matériels ou logiciels. Les techniques d'acquisition mobilisées pour un relevé dépendent d'un grand nombre de contraintes techniques ou contextuelles, et ce dans l'objectif de répondre aux impératifs idéalement exprimés dans un cahier des charges.

Ce guide met en avant les deux principales techniques de numérisation 3D utilisées de nos jours pour la reconstruction d'un modèle 3D sur base d'un modèle réel, soit par mesures directes (lasergrammétrie) soit par mesures indirectes (photogrammétrie). Ces techniques, bien que différentes, sont toutes deux basées sur l'extraction de données géométriques et/ou colorimétriques à l'aide de capteurs, actifs dans le cas des lasers, ou passifs dans le cas des photographies. S'il peut y avoir des contre-indications pour l'utilisation d'une de ces techniques en fonction de l'échelle, de la complexité, des matériaux et de l'environnement direct ou indirect de l'objet patrimonial à relever, il faut considérer ces techniques comme complémentaires. En effet, corollairement aux exigences en termes de résolution, précision, exhaustivité du relevé, l'utilisation de plusieurs techniques peut être requise. Une fois les données collectées, une succession plus ou moins importante d'étapes sera généralement nécessaire pour générer le modèle final qui pourra être décliné en différents livrables 2D ou 3D.

Comme pour le matériel, l'utilisation conjointe de différents logiciels complémentaires

permettant la restitution d'un modèle 3D peut s'avérer nécessaire. Il faut cependant rester vigilant sur la multiplication des étapes de traitement des données brutes pouvant altérer qualitativement (filtrage, interpolation) et quantitativement (discrétisation, décimation, échantillonnage) l'information initialement capitalisée. Cette problématique actuelle, inhérente à la traçabilité du processus dont résulte un modèle 3D implique une bonne connaissance des ressources matérielles et logicielles mobilisées pour sa création. Ce guide « Matériels et Logiciels » en complément des autres livrables du Consortium 3D peut servir de support pour évaluer ou rétro-analyser la validité de l'objet numérique livré à condition que le prestataire veille à conserver les métadonnées des fichiers sources et puisse renseigner les étapes intervenues pour produire les différentes versions (paradonnées, rapport, notice).

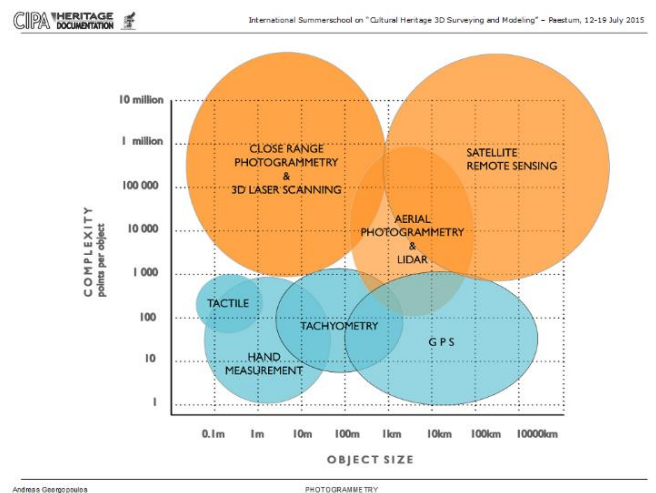


Figure 1 - Comparison of 3D Laser Scanning and Other 3D Measurement Techniques. (W. Boehler, 2006)

Le regroupement des différents partenaires du Consortium 3D a été l'occasion de partager l'expertise et le savoir-faire - tant théorique que pratique - de ses membres, sur les modalités de production d'un objet numérique patrimonial.

Un premier objectif a donc été de fournir une synthèse des outils et des technologies, d'offrir un point de vue sur les matériels, les logiciels et les pratiques employées le plus communément au sein de nos structures et disciplines respectives. Cet objectif a bien été mené sous réserve d'une non-exhaustivité tant il était utopique et fastidieux d'y référencer les spécificités propres à chaque membre. Une première étape a nécessité un travail de recensement et d'inventaire des solutions matérielles et logicielles utilisées au sein du Consortium 3D. Le choix a ensuite été fait de capitaliser les informations sur les méthodes et pratiques partagées, en accord avec l'état de l'art technologique et scientifique en matière de modélisation et numérisation 3D.

Le deuxième objectif a été d'inclure un apport critique sur les « Matériels et Logiciels » afin de restituer un retour sur expérience objectif dans l'utilisation de ces outils. C'est la raison pour laquelle, au-delà de la structuration et de l'énumération de spécificités techniques, ce livrable comporte des informations essentielles et des recommandations dans le cadre spécifique de la modélisation/numérisation 3D appliquée sur des objets patrimoniaux.

Le troisième objectif a été celui d'exercer une veille technologique primordiale dans un domaine qui évolue très rapidement. Ce travail, effectué sur une période quadriennale, nous a permis de prendre le recul et la distance critique nécessaire au sein d'une offre matérielle et logicielle en pleine expansion. Par conséquent, à défaut d'exhaustivité notre travail répertorie les outils offrant à la fois la stabilité et la maturité requises pour une utilisation courante au sein de notre communauté. En quatre ans seulement, de nombreuses technologies matérielles ou logicielles se sont en effet retrouvées désuètes (scanner à triangulation optique) ou marginales (photomodélisation) ; remplacées par d'autres nouvelles, innovantes et prometteuses. Cette évolution croissante et soutenue de l'ensemble de

ces ressources nous a alertés sur l'obsolescence probable et rapide de certaines informations (notamment dans la section « Logiciels » et ses annexes) reprises dans ce guide.

Nous espérons que ce guide « Logiciels et Matériels » ainsi que ses annexes offre une approche alliant didactique et précision technique nécessaire pour diffuser au sein de la sphère du patrimoine numérique les points théoriques essentiels pour une bonne pratique/utilisation des objets numériques patrimoniaux.

Notions de précision et d'incertitude ; géométrie, colorimétrie et métrologie

Cette problématique aurait pu faire l'objet d'un volet voire d'un atelier thématique dédié malgré la difficulté d'aborder ce sujet d'une grande complexité. Il convient tout d'abord de spécifier et d'accepter que malgré la haute précision des technologies actuelles, tout modèle 3D est par essence imparfait. Que l'on se réfère aux données brutes issues du capteur jusqu'à sa formalisation finale, un certain degré d'incertitude est inhérent à l'objet numérique. De nos jours, un modèle issu d'une numérisation 3D ne restitue que l'apparence visuelle de l'objet, selon les conditions d'éclairage d'un moment donné. Ces notions de précision et d'incertitude sont relatives au matériel mobilisé pour la phase d'acquisition et aux processus de traitements logiciels qui ont suivi, et qui peuvent altérer la qualité géométrique et/ou colorimétrique.