

Sandra Zanella, Jean-Pierre Brun, Martine Denoyelle, Pierre Rouillard et Stéphane Verger (dir.)

Les archives de fouilles : modes d'emploi

Collège de France

Entre droits et devoirs. L'archéologie face au défi de l'Open Data : le MOD, Mappa Open Data archaeological archive

Maria Letizia Gualandi

DOI : 10.4000/books.cdf.4905
Éditeur : Collège de France
Lieu d'édition : Paris
Année d'édition : 2017
Date de mise en ligne : 18 avril 2017
Collection : Institut des civilisations
ISBN électronique : 9782722604650

Ce document vous est offert par Collège de France



<http://books.openedition.org>

Référence électronique

GUALANDI, Maria Letizia. *Entre droits et devoirs. L'archéologie face au défi de l'Open Data : le MOD, Mappa Open Data archaeological archive* In : *Les archives de fouilles : modes d'emploi* [en ligne]. Paris : Collège de France, 2017 (généré le 28 avril 2017). Disponible sur Internet : <<http://books.openedition.org/cdf/4905>>. ISBN : 9782722604650. DOI : 10.4000/books.cdf.4905.

Ce document a été généré automatiquement le 28 avril 2017.

Entre droits et devoirs. L'archéologie face au défi de l'*Open Data* : le MOD, *Mappa Open Data archaeological archive*

Maria Letizia Gualandi

- 1 Depuis plus de vingt ans les archives de fouilles, rassemblant les données brutes (*raw data*), sont rédigées en format digital : fiches d'enregistrement stratigraphique et fiches de quantification de mobilier, tableaux de périodisation et matrix, plans, coupes et représentations graphiques de différents types, photographies et enregistrements video, SIG et ce qui est normalement appelé « l'écriture grise », c'est-à-dire les rapports de fouilles et les journaux de fouille. Il s'agit d'une importante quantité de données, enregistrée pendant le travail sur le terrain ou bien dans la suite immédiate des opérations, qui s'est considérablement accrue en raison de l'augmentation en Europe des pratiques liées à l'archéologie préventive et qui s'ajoute au très grand nombre documents de fouilles en format papier produits pendant les cent cinquante dernières années.
- 2 De toute la masse de données produite pendant une opération de fouille, seule une partie minimale est destinée à apparaître dans les publications de fouille, éditées principalement sur papier, où sont présentes, notamment, des données « interprétées », sans l'appui de la documentation brute nécessaire pour relire et réévaluer ces données.
- 3 En dépit de cela, cette documentation est, dans la plupart de cas, tout ce qui reste d'une mine d'informations que la fouille, qui est une pratique non reproductible, a contribué à détruire. Par conséquent la consultation des *raw data* est de facto le seul élément qui peut permettre une reproduction et une nouvelle analyse du processus interprétatif, indispensable point de départ pour pouvoir formuler des nouvelles hypothèses et reconstructions historiques. Ainsi, en groupant ces données de différentes façons, il est possible d'apporter des réponses à diverses questions, ou encore de les utiliser avec des


objectifs différents. Pour reprendre une expression du langage informatique, nous pouvons affirmer que les *raw data* sont les « code source » de l'archéologie.

- 4 Un autre facteur à prendre en considération est la date de publication des résultats de fouilles (quand ils sont publiés !). Leur parution est en général très éloignée, souvent de plusieurs années, de la fin du chantier de fouille et celui qui a produit ces données se garde bien de les rendre publiques par peur, pour ainsi dire, d'en perdre le copyright. Cette pratique, malheureusement très largement répandue, se base sur un malentendu entre « propriété » intellectuelle et « paternité » intellectuelle. Ceci est compréhensible, mais certainement pas justifiable car les deux notions sont bien différentes.
- 5 Avec « propriété » intellectuelle, on entend l'ensemble des droits exclusifs qui sont accordés sur les produits de l'esprit humain. Sur la base de ces principes, la loi donne aux titulaires de la propriété intellectuelle le monopole absolu dans la gestion et l'exploitation de leurs créations. Cependant, le développement des moyens de communication numériques a amené, au cours des dernières années, au réexamen des principes fondant la propriété intellectuelle. Une des principales explications à cette réflexion tient dans le détachement des créations intellectuelles de leur support matériel traditionnel, en particulier pour les livres imprimés. Dans ce contexte, des expériences innovantes telles que le copyleft ont vu le jour. Ce type de licence permet au détenteur du droit d'auteur de céder, par le biais d'une série de licences, certaines libertés aux utilisateurs de l'ouvrage : le créateur revendique alors la « paternité » intellectuelle plutôt que la propriété intellectuelle. Ce modèle, apparu et développé notamment dans le monde du logiciel informatique, s'est étendu, ces dernières années, à tous les œuvres produites par l'intellect, en parallèle aux mouvements Creative Commons, Open Access, Open Content etc.
- 6 Toute personne qui fait de la recherche sur le terrain sait que la publication scientifique d'une fouille nécessite un long temps d'élaboration, pendant lequel celui qui a produit les données de fouille doit pouvoir jouir d'un « droit de prélation » sur leur publication (un droit qui doit, dans tous les cas, être limité et défini dans le temps). Dans l'intervalle, les données vieillissent, alors que la recherche doit pouvoir s'alimenter de données « fraîches » et inédites. Il est donc indispensable que les *raw data* soient rapidement et facilement accessibles, avant même la publication scientifique de la fouille.
- 7 Aujourd'hui il est possible de protéger de manière efficace la paternité intellectuelle des données publiées dans les archives ouvertes grâce aux lois et conventions dont souvent les archéologues ignorent même l'existence. La plupart des chercheurs connaissent l'existence et la signification des codes ISBN (International Standard Book Number) et ISSN (International Standard Serial Number) (fig. 1) mais peu connaissent le code DOI (Digital Object Identifier), qui lie de façon pérenne un document à son auteur (ou ses auteurs), ou les licences Creative Commons (fig. 2) qui, à travers le monde, permettent le remploi créatif d'œuvres nées de l'esprit d'autrui, dans le respect des lois existantes dans chaque nation.

Fig. 1

ISBN
(*International Standard Book Number*):
numéro international qui permet d'identifier de manière unique chaque édition de chaque livre publié.

ISSN
(*International Standard Serial Number*):
numéro international qui permet d'identifier de manière unique une publication en série.



mappa

www.mappaproject.org

CODES ISBN ET ISSN.


© DR

Fig. 2

Digital Object Identifier (DOI):
mécanisme d'identification de ressources numériques, comme un film, un rapport, des articles scientifiques ou tout autre type d'objets.

10.4456/MAPPA.2013.21

Creative Commons (CC):
ensemble de licences régissant les conditions de réutilisation et/ou de distribution d'œuvres, notamment diffusées sur internet.

 www.mappa-project.org

Code DOI et licences Creative Commons.

© DR

- 8 Dans tous les cas, même après la publication de la fouille, la documentation – papier ou digitale – est peu utilisée et demeure dans des classeurs et DVD, eux-mêmes déposés dans les archives poussiéreuses des organes de tutelle, où ils sont, en théorie, consultables. Mais tout le monde sait que pour différentes raisons, il n'est pas toujours facile d'y accéder. Même si les fouilles les plus importantes finiront, un jour ou l'autre, par être publiées, il y a un nombre infini d'interventions « mineures » comme, par exemple, les enquêtes d'archéologie préventive d'extension limitée – tranchées pour la vérification de conduits souterrains, carottages préalables à l'édification d'une parcelle, etc. – dont on perd souvent complètement la trace. Leurs résultats ne seront jamais publiés, car considérés comme peu significatifs, fragmentaires et donc, difficilement explicables. Alors qu'ils pourraient mettre en lumière d'autres données, ou, au contraire, fournir la pièce manquante utile à la reconstruction d'un puzzle interprétatif de plus grande envergure.
- 9 Tout cela pénalise la recherche qui, elle, ne peut faire abstraction d'une libre et rapide circulation des données. Cela permettrait de dépasser le paradoxe pour lequel – comme l'a ironiquement souligné John Wilbanks, directeur exécutif du projet Science Commons – au moment historique où nous disposons de technologies qui rendent visibles les données scientifiques au niveau international et de systèmes de distribution qui permettent d'étendre la collaboration et d'accélérer le rythme et l'intensité des découvertes nous verrouillons l'accès aux données et nous empêchons l'utilisation de technologies d'avant-garde qui auraient un fort impact sur la diffusion des connaissances.

- 10 Mais, à bien regarder, sont également pénalisés l'activité de tutelle du patrimoine archéologique – qui appartient à tous et non à celui qui a fouillé ou bien à celui qui a la responsabilité de fouiller – et, par conséquent le développement d'une méthode de planification territoriale qui doit être respectueuse des restes ensevelis : comment est-il possible de penser d'autoriser ou d'interdire la réalisation d'une infrastructure en absence de données scientifiques complètes et à jour ?
- 11 Ces brèves remarques suffisent pour se rendre compte que, pour l'archéologie, l'ouverture des archives contenant les *raw data* des fouilles est une nécessité, plus encore un devoir (fig. 3) et que leur mise en ligne est un chemin aujourd'hui praticable, dans une optique d'archéologie durable, pour faire progresser les connaissances sans devoir supporter de coûts prohibitifs.

Fig. 3



La révolution des médias

© DR

- 12 Mais une Archéologie 2.0 apparaît encore comme une réalité bien lointaine, au moins en Italie, en dépit de la diffusion croissante de la philosophie *Open Data*. Au sentiment de méfiance (pour ne pas parler d'aversion) ressenti par beaucoup à l'égard de l'ouverture des archives *raw data*, il y a diverses raisons :
- l'incapacité à reconnaître à la publication des *raw data* la dignité de publication, qui est seulement conférée aux synthèses interprétatives. Si, en fait, nous considérons que l'activité de fouille est une activité de recherche qui produit des données uniques et irremplaçables et qu'une recherche respectable s'achève avec la publication des données, il semble évident que la publication des *raw data* et de la « littérature grise » doit être considéré comme étant une publication à part entière. De ce manque de reconnaissance,

vient la crainte de nombreux étudiants et chercheurs de perdre leur *copyright* sur les données produites par leur travail sur le terrain et dont nous avons déjà parlé ;

- la crainte d'exposer à l'opinion publique une documentation qui, par définition, est susceptible d'être soumise à des corrections dans ses éléments interprétatifs, mais qui est d'autant plus précieuse qu'elle contient des observations essentielles qui ne sont possibles uniquement qu'au moment où la stratigraphie est démontée et mise au rebus. Il a été dit avec raison qu'un *dataset* parfait est un *dataset* suspect ;
- la méfiance à l'égard du web – encore largement répandue vingt-cinq ans après la création, en 1991, du *World Wide Web* – et, par conséquent, envers la révolution qui a suivi des modèles de diffusion de la connaissance. Une révolution – il faut l'admettre – comparable à celle produite par l'introduction, en 1455, de l'imprimerie, qui eut elle aussi de nombreux adversaires (fig. 4).

Fig. 4



Avantages des archives *Open Data*

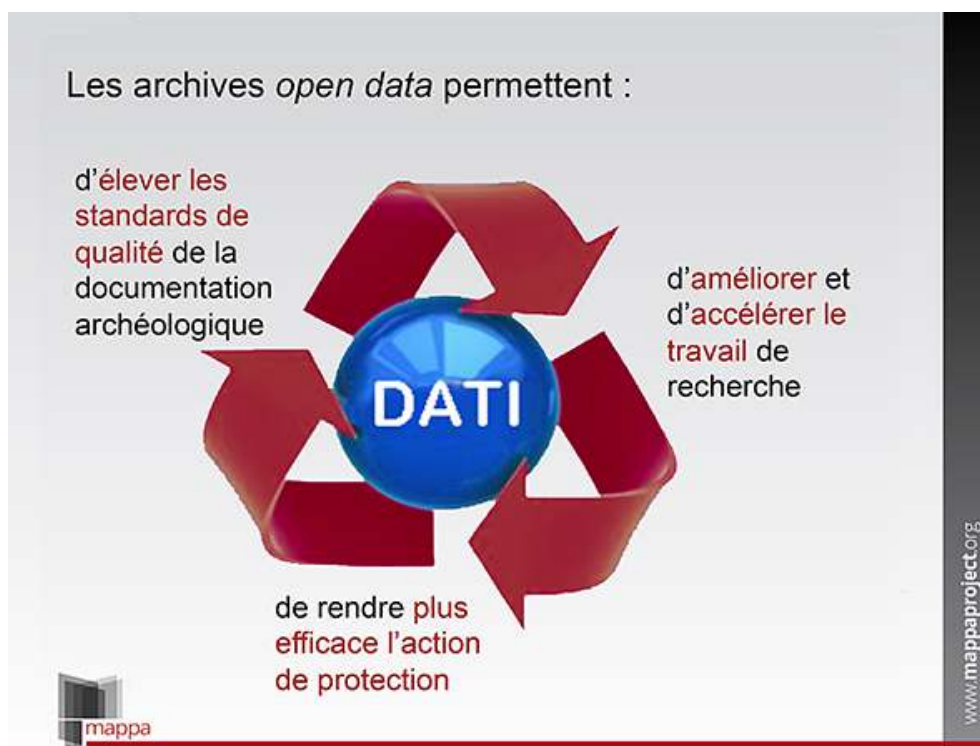
© DR

- 13 Un quart de siècle plus tard (un intervalle de temps similaire à celui qui nous sépare aujourd'hui de la naissance du *www*) d'illustres personnages, tels que, par exemple, l'abbé bénédictin Johannes Trithemius (1462-1516), continuaient à mépriser les livres imprimés. Cela parce qu'ils les jugeaient bien plus laids que les manuscrits et parce qu'il considérait le papier comme un matériel bien plus fragile et moins durable que le parchemin, mais surtout parce qu'ils craignaient que l'impression conduise à la prolifération incontrôlée des textes. Quiconque pourrait publier quoi que ce soit comme cela a été démontré par la diffusion en milliers d'exemplaires des 95 *Theses* de Luther, et avec toutes les conséquences que cela a impliqué pour l'histoire de l'humanité. Le duc Frédéric de Montefeltro avait même interdit l'entrée des livres imprimés dans la bibliothèque ducale

d'Urbino, atteignant le paradoxe de faire retranscrire à la main des volumes publiés sur papier !

- 14 En effet, chaque donnée peut être plus ou moins fiable indépendamment des modalités de publication et de la rapidité de la diffusion. Normalement les données archéologiques publiées dans la littérature scientifique ont un certain degré de fiabilité selon la crédibilité scientifique de son auteur et l'éventuel processus de *peer review* auquel la publication a été soumise. Avec les *Open Data* la crédibilité est déterminée par la possibilité de vérifier le processus de création de la donnée. La validation est, donc, confiée à la responsabilité individuelle de chaque archéologue, dans une optique de révision entre pairs (*open peer review*).
- 15 Il est donc évident que l'ouverture des données est essentiellement aujourd'hui une question culturelle : cela signifie, tout d'abord, de modifier notre manière de nous rapporter à l'archéologie ; de prendre réellement conscience des potentialités offertes par les données ouvertes. En fait, les archives *Open Data* sont encore à découvrir. Pour cela, il est nécessaire que les archéologues voient dans le partage des informations le vrai moteur de la connaissance et du développement de la recherche, de la protection du patrimoine et d'un nouveau et plus large concept de culture archéologique ; il est donc nécessaire, en d'autres termes, que les archéologues prennent conscience du fait que le recyclage des données est possible grâce à l'*Open Data* (fig. 5).

Fig. 5



MOD : page d'introduction (<http://www.mappaproject.org/archivio-digitale>)

© DR

- 16 Il permet :
- d'élever les standards qualitatifs de l'enregistrement archéologique, grâce à la comparaison entre les systèmes de documentation ;

- d'améliorer et, surtout, d'accélérer le travail de recherche, parce qu'il permet de repérer avec facilité les données (on pense en particulier aux courts délais imposés par les enquêtes de l'archéologie préventive) ;
- d'augmenter l'efficacité de la protection, car dans une communauté éduquée avec l'idée d'un patrimoine historique et archéologique en tant que bien commun, un ample partage des informations devient progressivement une garantie de contrôle, de qualité et de conservation ;
- d'utiliser les données à de diverses fins : tourisme, planification et beaucoup d'autres qu'il est, à l'heure actuelle, difficile d'imaginer...

Le mod (*mappa Open Data archaeological archive*)

- 17 À ce débat l'Université de Pise a offert sa propre contribution en mettant à disposition de tous le MOD (*MAPPA Open Data archive*), la première archive archéologique italienne en *Open Data*¹. Développé dans le cadre d'un projet de recherche intitulé MAPPA – Metodologie Applicate alla Predittività del Potenziale Archeologico. L'area urbana di Pisa² –, financé par la Région Toscane et par l'Université de Pise, le MOD a été mise en ligne dans sa version bêta en 2012, quand en Italie le débat sur les *Open Data* en était à ses origines.
- 18 Le MOD a été le fer de lance dans le difficile processus d'adaptation de l'archéologie italienne à des standards comme ceux de l'anglais ADS, *Archaeological Data Service*³, qui depuis le 1996 fournit un extraordinaire support à la recherche archéologique en mettant à disposition de tous – archéologues mais pas seulement – des archives *open* scientifiquement fiables et en se faisant de promoteur des bonnes pratiques des données digitales.
- 19 Depuis le 2013 le MOD est *on line* dans sa version complète (fig. 6)⁴.

Fig. 6

italiano | english

MAPPAAopenDATA

Conservare, disseminare, collaborare. L'archeologia libera e aperta è la nuova frontiera della ricerca.

Il MOD (MAPPAA Open Data) è l'archivio digitale archeologico pensato all'interno del gruppo di lavoro del Progetto MAPPA per conservare e disseminare la documentazione archeografica (Dataset) e la letteratura grigia (Relazioni) prodotta nel corso di una qualsiasi indagine archeologica.

Il Progetto MAPPA attraverso il MOD vuole tutelare chi ha prodotto la documentazione in tutte le sue forme sia a livello di dato grezzo archeografico (Dataset), sia a livello di letteratura grigia (Relazioni), per questo ha individuato nell'apposizione di un DOI ad ogni Dataset e/o Relazione e nell'utilizzo della licenza CC BY-SA 3.0, il modo per consentire contemporaneamente la corretta circolazione dei dati e la tutela della paternità intellettuale.

Il MOD nasce dall'esperienza di Pisa, ma con la collaborazione di tutti ambisce a diventare una piattaforma condivisa per tutto il territorio nazionale. Chiunque voglia collaborare aggiungendo nuovi Dati all'archivio può contattarci all'indirizzo info@mappaproject.org

L'inserimento della documentazione (dataset/relazioni) di un intervento all'interno del MOD è fatta su base volontaria dai detentori della paternità intellettuale dei dati (vedi Disclaimer). L'immissione di tali dati nel sistema è curata dal Progetto MAPPA.

L'accesso al MOD è gratuito.

L'accesso all'applicazione presuppone l'accettazione delle **Condizioni di Servizio** e del **Disclaimer**.

>> ACCEDI AL MOD

www.mappaproject.org

MOD : protection du droit de paternité intellectuelle des données

© DR

- 20 Pensé en utilisant comme premier set les données relatives aux opérations archéologiques effectuées dans l'aire urbaine de Pise, le MOD a commencé progressivement à accueillir également les *raw data* de recherches archéologiques qui ont eu lieu en dehors de Pise. Depuis le début, en effet, l'objectif n'était pas de réaliser un prototype qui aurait pu servir de modèle pour beaucoup d'archives différentes, peu compatibles les unes avec les autres, mais plutôt de proposer le MOD même comme l'archive italienne des données archéologiques. Un lieu virtuel et une structure digitale dans laquelle peuvent enfin trouver place toutes les données archéologiques généralement inaccessibles sur le réseau, comme c'est le cas dans l'ADS de York. En temps de crise économique et avec des financements pour la recherche réduits au plus bas, l'idée a été celle de mettre à disposition de la collectivité l'important travail déjà existant (et déjà financé) dans le cadre du projet MAPPA.
- 21 L'insertion dans le MOD de la documentation produite dans le cadre d'une fouille archéologique équivaut en tous points à une publication (fig. 7), dont la paternité intellectuelle est sous la tutelle du code DOI. Son usage est consenti sur la base des licences *Creative Commons* CCBY (c'est-à-dire avec obligation de citation de l'auteur, comme c'est le cas pour les ouvrages imprimés) ou CCBYSA (avec obligation de citation de l'auteur et de partage sous le même format *open* des résultats obtenus grâce à l'utilisation de ces données)⁵.

Fig. 7

mod
mappaopendata

Protection du droit de paternité intellectuelle

À chaque document est associé un code DOI (*Digital Object Identifier*), l'équivalent numérique de l'ISBN et de l'ISSN, reconnu internationalement

10.4456/MAPPA.2013.21

Chaque auteur choisit parmi deux licences *Creative Commons*:

- *Creative Commons BY (attribution)*: l'œuvre peut être librement utilisée, à la condition de l'attribuer à l'auteur en citant son nom;
- *Creative Commons BY SA (share alike)*: les œuvres dites dérivées doivent être proposées au public avec les mêmes libertés (sous les mêmes options *Creative Commons*) que l'œuvre originale.

mappa

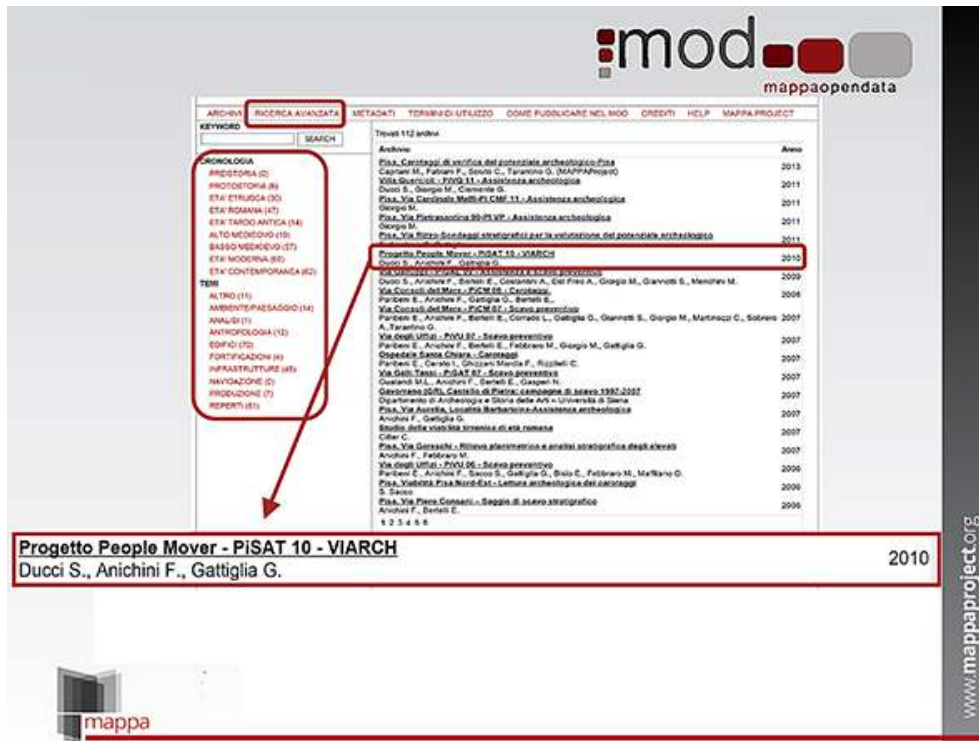
www.mappaoproject.org

MOD : page d'identification des interventions

© DR

- 22 Une publication en temps rapides, à coût zéro et, du moment qu'elle est *on-line*, avec un pouvoir de dissémination qu'aucune publication papier pourra jamais espérer avoir, notamment en ces moments de crise, dans lesquels il est toujours plus difficile de trouver des fonds pour publier et acheter des livres⁶.
- 23 Dans le MOD toutes les interventions archéologiques, de tout type ou importance – de la grande fouille à la mise en œuvre de conduites – ont la même visibilité. Chaque intervention est identifiée avec le nom de la localité où elle a été réalisée, la typologie, l'année d'exécution et l'auteur ou les auteurs (fig. 8).

Fig. 8



MOD : d'introduction de chaque intervention, avec une fiche synthétique de présentation
 © DR

- 24 Une division en macro-périodes historiques et thématiques archéologiques permet une première grande classification des données, qui guide l'utilisateur dans la navigation, alors que le champ « recherche avancée » permet d'affiner la consultation, en ajoutant aux champs thématiques et chronologiques l'année d'intervention, les mots connus à l'intérieur du nom du fichier, le nom de l'auteur, le format du fichier à rechercher.
- 25 Pour chaque intervention sont générées :
 - une page d'introduction (fig. 9), avec une fiche synthétique de présentation ;
 - une page dédiée à la « littérature grise » (fig. 10), où il est possible de télécharger les documents, avec les indications de la modalité exacte des renvois des documents ;
 - une page où il est possible de télécharger toute la documentation (*dataset*) (fig. 11), organisée par typologie : documentation graphique, photographique, écrite.

Fig. 9

The screenshot shows the MOD (Mappa Open Data) website interface. At the top, there is a navigation menu with links: ARCHIVI, RICERCA AVANZATA, METADATI, TERMINI DI UTILIZZO, COME PUBBLICARE NEL MOD, CREDITI, HELP, and MAPPA PROJECT. The main header features the 'mod' logo and 'mappaopen data' text.

The page title is 'Progetto People Mover - PISAT 10 - VIARCH' by Ducci S., Anichini F., Gattiglia G., 2010. Below the title, there are two red-bordered boxes: 'Introduzione Relazione Dataset' and 'Introduzione'. An arrow points from the first box to the second.

The 'Introduzione Relazione Dataset' box contains the following information:

- Contatto principale:** Gabriele Gattiglia, Studio Associato InArcheo, g.gattiglia@inarcheo.com
- Identificatori:** ID: 00015, DOI RELAZIONE: 10.4456/MAPPA.2012.12, DOI DATASET: 10.4456/MAPPA.2012.13

The 'Introduzione' box contains a satellite map image and the following text:

Esecutore: Studio Associato InArcheo;
 Direttore Scientifico: Silvia Ducci (SBAT)

La VIARCH realizzata nel 2010 per il progetto People Mover, relativo all'esecuzione di un nuovo tracciato per la navetta tra l'aeroporto e la stazione ferroviaria di Pisa Centrale, ha previsto l'esecuzione di 4 carteggi (via Quarantola e via Goletta) e 3 ricognizioni (via Goletta, via San Giusto, via S. Agostino).

Below the map, there is a Creative Commons license icon (CC BY) and the text 'Questo archivio è rilasciato con licenza CC BY'. At the bottom of the page, there are logos for 'UNIVERSITÀ DI PISA', 'REGIONE TOSCANA', and 'DAS'.

The footer includes the 'mappa' logo and the website address 'www.mappaproject.org' on the right side.

MOD : page présentant le générique de l'opération
 © DR

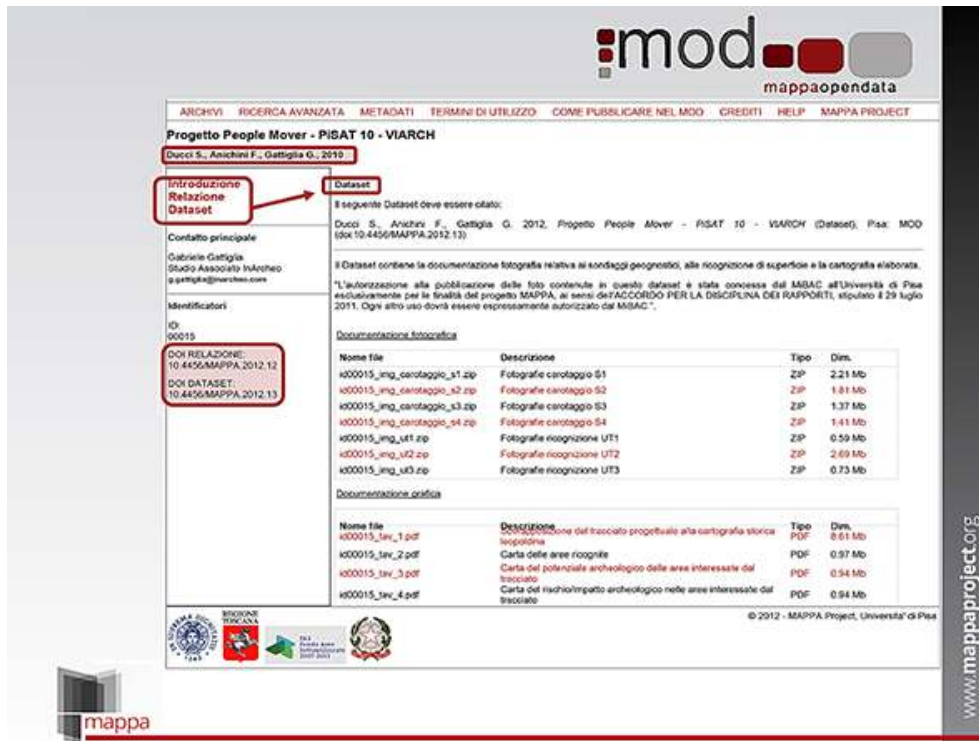
Fig. 10



MOD : page dédiée à la documentation (dataset), organisée par typologie (documentation graphique, photographique, écrite)

© DR

Fig. 11



?

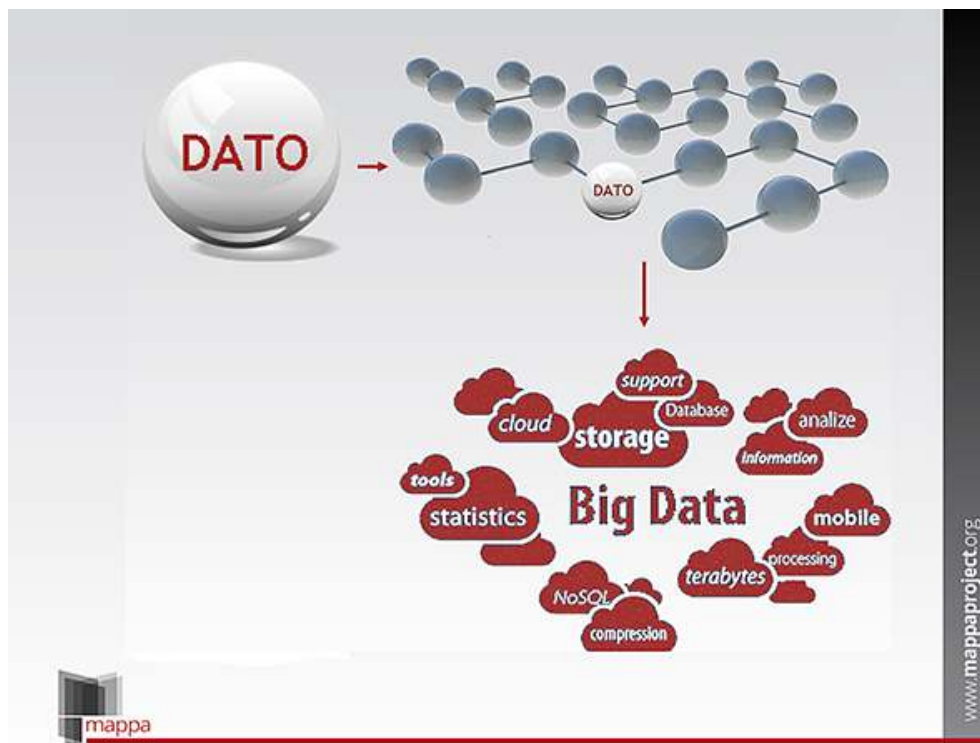
© DR

- 26 Dans chaque page sont toujours affichés le nom de l'auteur (ou des auteurs), le code DOI et la licence *Creative Commons* qui garantissent la reconnaissance de la documentation insérée dans le MOD comme étant une publication scientifique.
- 27 Puisqu'il s'agit d'une nouveauté absolue en Italie, dans le champ de la conservation et diffusion des données de fouille, le MOD a comme objectif de sensibiliser la communauté scientifique aux thématiques de l'*Open Data*, dans le propos de faire comprendre les grandes opportunités offertes par ce nouveau moyen de dissémination des données. Pour cela, le MOD a été conçu pour accueillir des documents créés à l'origine sous le format dwg, pdf, jpg, tiff, csv, xml, xls ou bien de simples documents scannés.
- 28 Cela suffit-il ? Non. Nous sommes conscients que les formats ne sont pas tous *open source* ou *machine readable* et que donc tous les documents n'atteignent pas les standards définis par la *Open Knowledge Foundation* en 2010, selon laquelle « *Open means anyone can freely access, use, modify, and share for any purpose (subject, at most, to requirements that preserve provenance and openness)* »⁷.
- 29 Mais nous avons décidé de nous en contenter, pour ainsi dire, dans le but de faire circuler le plus de matériel possible, en étant conscients qu'un travail d'optimisation des formats pourra être fait dans une étape ultérieure. Pour l'archéologie italienne, en effet, il était nécessaire de débiter, en utilisant le matériel documentaire disponible, sans prétention de tout « réécrire ». En d'autres termes, il était indispensable de faire le premier pas, le plus important : changer la mentalité des archéologues et vaincre leur méfiance pour qu'ils puissent découvrir l'utilité et les énormes possibilités que le *sharing* peut offrir, en

s'engageant conjointement sur une confrontation des modalités et standards nécessaires pour rendre les données plus utiles pour le travail de tous.

- 30 Chaque donnée, en effet, a son propre potentiel d'information, mais, si on met ensemble plusieurs données en le croisant avec d'autres (fig. 12), le potentiel informatif grandit de manière exponentielle.

Fig. 12



De la donnée isolée au big data

©DR

- 31 Mettre à disposition des chercheurs d'importantes quantités de données, qui peuvent être traités automatiquement, signifie amener peu à peu l'archéologie dans le champ des *big data* et espérer pouvoir arriver à des analyses sur une ample échelle, effectuées avec des méthodes statistiques et des modèles mathématiques⁸ dont l'application est commune dans tous les autres champs de la science. Mais aujourd'hui, cela est encore impensable en archéologie, notamment en raison de la manière fragmentaire dont les données sont disséminées.
- 32 Il est évident que le MOD est un instrument encore largement à améliorer et, en dépit du bon accueil dont il a fait l'objet par nombreux archéologues italiens, la route à parcourir pour abattre le mur de la méfiance qui entoure un instrument de recherche si innovant est encore longue. À cela s'ajoutent les doutes sur l'utilité de publier les *raw data* en l'absence d'une synthèse interprétative, notamment si elles sont mises à disposition d'usager non-spécialistes et donc destinées à se perdre au milieu d'une accumulation immense de données. C'est en effet un problème réel, dont la portée, à mon avis, devrait néanmoins être évaluée après une adéquate période d'expérimentation des archives *Open Data* et un monitoring attentif des accès et des réemplois effectifs des informations. L'expérience de l'*Archaeological Data Service* est éclairante en ce sens. Il suffit de regarder

son histoire, commencée il y a vingt ans, pour se rendre compte des transformations qu'il a subies en fonction d'exigences de recherche, tutelle et didactique toujours renouvelées, qui sont venues se matérialiser au long des années⁹. Le risque est celui de perdre temps, énergies et ressources en études de faisabilité et discussion théoriques infinies qui souvent ne portent à aucun résultat concret.

- 33 N'importe quelle archive, *Open Data* ou pas, est en effet censée accueillir et conserver « toutes » les données disponibles. Sélectionner en amont ce qu'il est opportun de rendre public et ce qui ne l'est pas, en privant l'utilisateur de choisir librement les informations à utiliser pour ses propres objectifs, est un arbitrage et une prévarication des facultés de chacun de choisir les informations qu'il considère plus appropriées pour ses propres fins. Il est évident que l'utilisateur doit être accompagné dans la navigation parmi une telle masse de données : il est nécessaire de choisir le format plus adapté pour chaque typologie d'information, « catégoriser » les données, utiliser des métadonnées. Mais, comme le savent ceux qui ont déjà fait l'expérience de rechercher au moins une fois quelques documents dans une archive traditionnelle, le charme des archives est justement dans le fait que, s'il n'est pas souvent facile de trouver ce que l'on cherche, en revanche il est très facile de faire des découvertes extraordinaires en se baladant parmi les boîtes et classeurs.
- 34 En ayant cela à l'esprit, après la mise *on line* du MOD, dans le laboratoire MAPPA de l'Université de Pise on a commencé à travailler dans trois directions différentes :
- l'étude du format des données, pour faciliter le emploi, et de la structure des métadonnées, pour améliorer la visibilité et faciliter l'accès ;
 - la création d'une Open School of Archaeological Data (fig. 13)¹⁰, dans laquelle sont fournies aux archéologues les connaissances techniques de *data-archaeology*, qui leur permettent d'apprendre à travailler avec les données ouvertes, à les manipuler, à produire des visualisations et notamment à en connaître les potentialités. L'objectif est celui d'encourager l'emploi des archives *Open Data* et d'en faire croître la demande par le bas ;
 - la création d'une collection de *Data-book* associée au MOD (fig. 14)¹¹, qui associent un niveau de synthèse interprétative, représenté par un livre de type traditionnel, à un niveau plus innovant, qui consiste dans la publication des *raw data* en format *open*. Chaque *Data-book* est muni d'un code ISBN, comme toute publication imprimée, et il est diffusé à titre gratuit en format digital *open acces* sur le site www.mappaproject.org, ou bien il peut être acheté auprès de la maison d'édition en format papier en modalité *Print on Demand* (donc à un prix contenu). Les données de fouilles, qui sont à la base de chaque article imprimé dans le *Data-book*, sont publiées en tant que *Open Data* dans le MOD¹² et elles sont soumises à une licence *Creative Commons*, identifiées par un code *DOI*, qui en garanti la paternité intellectuelle. N'importe qui peut publier les résultats de ses propres recherches sur le terrain dans les *Data-book*, à condition d'une acceptation par les *peer reviewers*.

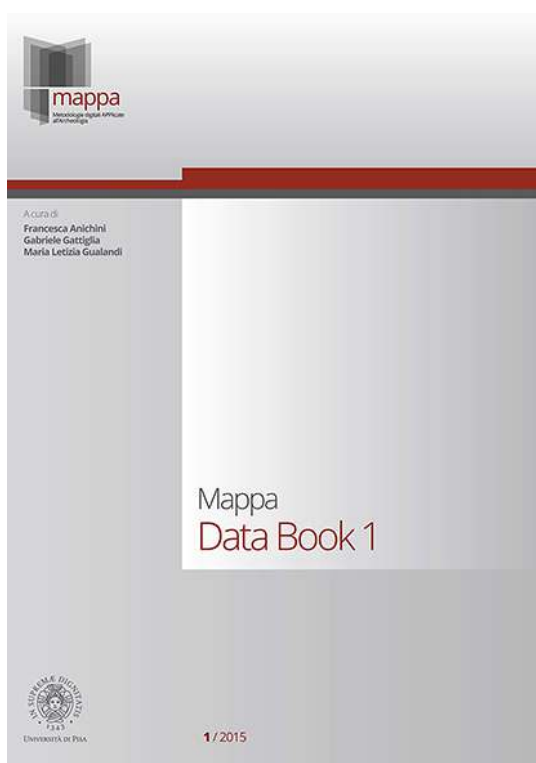
Fig. 13



OSAD – Open School of Archaeological Data de l'Université de Pise (<http://www.mappaproject.org/open-school-of-archaeological-data>)

© DR

Fig. 14



Couverture du premier Data Book (http://www.mappaproject.org/wp-content/uploads/2011/08/ DATA_Book_completo_copertina.pdf)

© DR

- 35 Nous espérons que cette entreprise fera naître des stimuli concrets pour améliorer les standards de notre travail et pour nous améliorer, en mettant à profit ce que la technologie moderne nous offre de mieux pour dépasser les difficultés qui se sont créées dans ces années de crise.
- 36 Comme l'écrivait il y a désormais un siècle Albert Einstein¹³ :
- 37 « Ne prétendons pas que les choses vont changer si nous continuons à faire la même chose. La crise est la plus grande bénédiction qui puisse arriver aux gens et les pays parce que la crise apporte des progrès. La créativité naît de l'angoisse comme le jour vient de la nuit noire. C'est dans la crise que l'invention est née, les découvertes et les grandes

stratégies. [...] La vraie crise est la crise de l'incompétence. Le problème des personnes et des pays sont la paresse pour trouver les sorties et les solutions. Sans crise il n'y a pas de défi, sans défis la vie est une routine, une lente agonie. Sans crise il n'y a pas de mérite. C'est dans la crise où chacun doit donner le meilleur de soi-même, parce que sans crise tout vent est caresse. [...] Seulement discuter de la crise est la promouvoir et rester silencieux vis-à-vis de la crise est exalter le conformisme. Au lieu de cela travaillons fort. Finissons-en une fois avec les menaces vis-à-vis de la crise. C'est ces mêmes menaces qui forment la tragédie en empêchant les personnes de lutter et la surpasser par leur audace ».

BIBLIOGRAPHIE

Anichini, Dubbini, Fabiani, Gattiglia, Gualandi 2013 : F. Anichini, N. Dubbini, F. Fabiani, G. Gattiglia, M.L. Gualandi, *MAPPa. Metodologie Applicate alla Predittività del Potenziale Archeologico*, 2, Roma, Edizioni Nuova Cultura, 2013.

Anichini, Gattiglia, Gualandi 2015 : F. Anichini, G. Gattiglia, M.L. Gualandi, *MAPPa Data Book*, 1, Roma, Edizioni Nuova Cultura, 2015.

Anichini, Dubbini, Fabiani, Gattiglia, Gualandi 2015 : F. Anichini, N. Dubbini, F. Fabiani, G. Gattiglia, M.L. Gualandi, « Dall'Open Data alla predittività. Nuovi modi di far crescere l'archeologia italiana », *Archeologia e calcolatori*, supplemento 7, 2015, 121-135.

Anichini, Ciurcina, Fabiani, Gattiglia, Gualandi 2016: F. Anichini, M. Ciurcina, F. Fabiani, G. Gattiglia, M.L. Gualandi, « Tra diritti e doveri. L'Open Data nell'archeologia italiana », a cura di F. Stanco e G. Gallo in *Proceedings of Archeofoss. Free, Libre and open source software e open format nei processi di ricerca archeologica* (Catania 2013), Archaeopress, Oxford 2016.

Einstein 1913: A. Einstein, *Il mondo come io lo vedo*, 1931 [1934].

NOTES

1. http://mappaproject.arch.unipi.it/?page_id=454. Il MOD è publié dans : Anichini, Dubbini, Fabiani, Gattiglia, Gualandi 2013 ; le volume est téléchargeable depuis la page : <http://mappaproject.arch.unipi.it/wp-content/uploads/2013/07/MAPPa-VOL.2.pdf>. Il MOD est listé parmi les repositories conseillés par le « Journal of Open Archaeological Data », <http://openarchaeologydata.metajnl.com>.

2. www.mappaproject.org.

3. <http://archaeologydataservice.ac.uk>

4. <http://www.mappaproject.org/archivio-digitale>.

5. Anichini, Ciurcina, Fabiani, Gattiglia, Gualandi 2016.

6. Pour publier les « données brutes » et la « littérature grise » issus de n'importe quelle fouilles ou survey, partout dans le monde, il suffit de se mettre en contact avec le Laboratoire MAPPa, tel. +39 (0) 50-2215817, mail info@mappaproject.org ; ou bien directement avec l'auteur : Dipartimento di Civiltà e forme del sapere,

Università di Pisa, via dei Mille 19, 56126 Pisa,
tel. +39 050 2215662, mail : letizia.gualandi@unipi.it.

7. <http://opendefinition.org>.

8. Anichini, Dubbini, Fabiani, Gattiglia, Gualandi 2015.

9. <http://archaeologydataservice.ac.uk/about>.

10. <http://www.mappaproject.org/open-school-of-archaeological-data>.

11. Anichini, Gattiglia, Gualandi 2015 ; <http://www.mappaproject.org/mappa-data-note-1>.

12. http://www.mappaproject.org/?page_id=454.

13. Einstein 1931 page ?

AUTEUR

MARIA LETIZIA GUALANDI

Professeur au Dipartimento di Civiltà e Forme del Sapere de l'Università de Pise, Italie