



HAL
open science

Les études sur les mondes musulmans et les humanités numériques. Introduction

Antoine Perrier, Clément Salah

► **To cite this version:**

Antoine Perrier, Clément Salah. Les études sur les mondes musulmans et les humanités numériques. Introduction. *Revue des Mondes Musulmans et de la Méditerranée*, 2024, Les études sur les mondes musulmans et les humanités numériques, 153, 10.4000/130fs . hal-04872095

HAL Id: hal-04872095

<https://hal.science/hal-04872095v1>

Submitted on 7 Jan 2025

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Pre-print de Antoine Perrier, Clément Salah « Les études sur les mondes musulmans et les humanités numériques », *Revue des mondes musulmans et de la Méditerranée*, vol. 156, n°2, 2024, p. 9-24.

Lien en openaccess : <https://journals.openedition.org/remmm/21634>

Les études sur les mondes musulmans et les humanités numériques. Introduction

L'usage de méthodes et de pratiques liées au développement et à l'utilisation d'outils numériques (couramment appelé « humanités numériques » ou, en anglais, *Digital Humanities*¹) – qu'il s'agisse des techniques relatives à la numérisation de ressources documentaires, aux bases de données en ligne, à l'indexation des ressources disponibles sur le web, à la collecte automatique des données, à l'extraction et à l'enrichissement de textes, etc. – bouleverse la production du savoir scientifique². Les études sur les mondes musulmans ne sont pas en reste : la multiplication des corpus documentaires en ligne, l'accès à des millions de livres numérisés en arabe, mais aussi en persan, syriaque, langues berbères et turques, et le nouveau terrain que représente internet lui-même ont ouvert de nombreuses possibilités. Ce numéro de la *Revue des mondes musulmans et de la Méditerranée* propose une première exploration méthodique des conséquences de l'intégration des humanités numériques dans le champ des études francophones portant sur les mondes musulmans, sans prétendre à l'exhaustivité. En situant cette réflexion à plusieurs niveaux, ce dossier thématique explore tant la pratique des humanités numériques au sein des mondes musulmans que les changements qu'elles induisent dans la production d'un savoir scientifique sur la région, en particulier en France. Mobilisant historiens, historiens de l'art, archéologues, sociologues, anthropologues, ethnographes et archivistes, les sept contributions de ce dossier fournissent les premiers éléments d'un panorama des recherches actuellement menées au moyen des humanités numériques dans le contexte spécifique des études sur les mondes musulmans.

Ce contexte n'est pas toujours déterminant et ce champ d'études est traversé d'interrogations plus générales qui concernent toutes les sciences humaines et sociales. Il reste toutefois un ensemble de particularités touchant principalement aux langues de la région (écrites en alphabet non latin et présentes moins massivement que les langues occidentales sur internet),

1 La définition même des humanités numériques est objet de débats. Pour une vue d'ensemble de ces débats, voir Gold, 2012.

2 Les auteurs remercient Mercedes Volait pour sa précieuse relecture et ses encouragements constants pour la naissance de ce numéro.

aux fortes inégalités d'accès au numérique dans celle-ci, ainsi qu'aux difficultés croissantes pour se rendre sur le terrain, qui confèrent parfois une importance supplémentaire aux ressources en ligne qui peuvent suppléer, faute de les égaler, les enquêtes menées sur place. Les contributions de ce dossier en donnent un exemple et mettent en relief les méthodologies innovantes et les ressources mobilisées par les chercheurs de différentes disciplines. Loin de se limiter à un inventaire purement technique de nouveaux outils, ce dossier propose une réflexion sur les conséquences de la rupture numérique dans les méthodes et les objets des sciences humaines et sociales pratiquées dans la région. Cette introduction voudrait offrir une rapide présentation du cadre institutionnel et de la littérature déjà consacrée aux humanités numériques dans les mondes musulmans.

Dans le monde anglophone, une première réflexion méthodologique

Nous entendons ici les humanités numériques comme l'ensemble des méthodes et pratiques allant de la création de site web ou de bases de données à la numérisation de données (textes, mais également images ou sons), leur traitement par des outils informatiques (extraction de texte, lexicométrie, etc.), en passant par la représentation graphique de réseaux, le suivi de la réutilisation des textes ou encore la récupération et l'analyse des métadonnées produites sur internet. Le point commun de ces approches est qu'elles s'appuient toutes sur des outils numériques. L'objet de cet état de l'art n'est pas de donner un aperçu exhaustif des recherches anciennement ou actuellement menées par le recours aux humanités numériques, d'autant plus qu'un livre blanc en a déjà dressé un premier bilan il y a quelques années (Volait et Lucas, 2020), mais plutôt de souligner les réflexions des chercheurs inscrits dans ce domaine quant à l'intégration de ces outils et les changements qu'ils induisent dans le travail de recherche sur les mondes musulmans.

Les chercheuses et les chercheurs anglophones sont les premiers à avoir abordé le rapport entre les humanités numériques et les études sur les mondes musulmans sous un angle réflexif. Si les réflexions sur l'usage des humanités numériques sont plus ou moins anciennes, ce n'est qu'à la fin des années 2010 qu'une série d'initiatives ont émergé au profit d'un questionnement plus systématique. En 2017, Elias Muhanna, professeur de littérature arabe classique et d'histoire médiévale à l'université de Brown (Rhode Island, États-Unis), a dirigé une première synthèse en anglais (Muhanna, 2017). Rassemblant dix articles, *The Digital Humanities and Islamic & Middle East Studies* aborde quelques grands thèmes des humanités numériques, notamment les questions relatives à la numérisation, la visualisation, la fouille de texte, les bases de données, la cartographie ou encore l'e-publication. En se voulant relativement accessible pour un public de non-spécialistes, celui de chercheurs n'utilisant pas nécessairement les outils numériques ou ceux qui les mobilisent mais ne travaillent pas sur les mondes musulmans, l'ouvrage souligne, tant dans sa forme que dans ses conclusions, que les recherches menées au moyen des humanités numériques ne nécessitent pas impérativement de

connaissances poussées en informatique. Les contributions rassemblées dans *Digital Humanities and Islamic & Middle East Studies* rappellent que celles-ci commencent, pour leur majorité, avec des logiciels dont l'usage est déjà très répandu. Enfin, les auteurs notent que, contrairement à certains présupposés, les humanités numériques n'offrent pas de solutions à toutes les questions scientifiques et ne sauraient remplacer l'intervention humaine. Ils apportent enfin la démonstration que le Moyen-Orient est tout aussi concerné que le reste du monde par les transformations numériques.

De façon presque concomitante de la publication de la synthèse dirigée par E. Muhanna, Cornelis van Lit, historien médiéviste à l'université d'Utrecht (Pays-Bas), a transformé ce qui était à l'origine un blog en revue scientifique en ligne consacrée à l'utilisation des technologies informatiques dans les domaines de la recherche qualifiés en anglais, aujourd'hui ou dans le passé, d'*oriental* : *The Digital Orientalist* a ainsi vu le jour en 2018³. Constatant que les défis technologiques et institutionnels sont communs à tous ces domaines, *The Digital Orientalist* vise à rassembler tant des spécialistes des mondes musulmans (travaillant sur le Maghreb, le Moyen-Orient, l'Afrique ou l'Asie) que des spécialistes d'autres sociétés et cultures asiatiques (de la Chine, de l'Inde, du Népal ou encore du Japon). La revue rejoint ainsi des dynamiques transrégionales qui se retrouvent dans d'autres spécialités, en particulier avec *The Digital Classicist*⁴, *The Digital Medievalist*⁵ ou encore *The Digital Humanist Initiative*⁶.

Outre la publication d'une première synthèse et la création d'une revue consacrée aux humanités numériques appliquées aux études sur les mondes musulmans, des réseaux de chercheurs, notamment en histoire médiévale ou en islamologie, se sont organisés aux États-Unis et en Europe. Dans ce cadre, l'*Islamicate Digital Humanities Network* (IDHN)⁷, créé en 2019, se donnait pour but premier de fédérer une communauté de chercheurs intéressés par l'exploration et la création d'outils numériques ou de techniques d'analyse informatique dans le domaine de l'histoire, de la littérature ou du droit islamique. Il organise dans cet objectif divers ateliers (*conferences*) en ligne, parfois généraux, parfois consacrés à des thèmes précis, comme celui consacré aux études sur le hadith en janvier 2021⁸, poursuivi en janvier 2023 par

3 <https://digitalorientalist.com/about/>. Voir également van Lit, Morris et Cevik, 2020.

4 <https://www.digitalclassicist.org/>

5 <https://journal.digitalmedievalist.org/>

6 <https://caiml.org/dighum/>. Voir également Werthner et al., 2024.

7 <https://idhn.org/>

8 <https://www.youtube.com/watch?v=vxVJHi3Ckm8>

celui sur les études sur le hadith et le droit chiite⁹. En outre, l'IDHN participe à la diffusion d'information et au partage d'une expertise interdisciplinaire en tenant à jour une liste d'outils, de bases de données et d'autres ressources. Il se propose de remplir un rôle de veille dans un contexte où les technologies évoluent très rapidement, tout en offrant un lieu de discussion et de réflexion critique. Son site internet propose enfin une série de ressources, assez limitées, ainsi que l'enregistrement de l'ensemble des ateliers organisés.

En parallèle de ces réflexions collectives et générales, certains chercheurs ont proposé des analyses plus circonstanciées autour d'une méthode, d'une pratique ou d'une implication de l'intégration des humanités numériques dans les recherches sur les mondes musulmans. De façon schématique, quatre grands thèmes ont occupé les chercheurs anglophones : (1) la numérisation ; (2) les bases de données et leur visualisation ; (3) les principes de réutilisation, de réemploi et de lecture distante (*remote reading*) des textes¹⁰ ; (4) internet comme objet d'étude.

Les spécialistes ont d'abord scruté les bouleversements induits par l'usage d'outils numériques dans la lecture et l'analyse des sources, notamment manuscrites. Les manuscrits font en effet l'objet d'un grand nombre de projets aboutissant à la publication de vastes corpus de textes, exploités par toutes les disciplines concernées (codicologie, histoire, histoire de l'art, littérature, philosophie...). L'analyse et la reconnaissance automatique des documents manuscrits (ci-après HTR, pour *Handwritten Text Recognition*) sont ainsi devenues la première étape presque obligatoire de nombreux projets qui bénéficient de la politique de numérisation massive des institutions patrimoniales dans le monde occidental depuis une vingtaine d'années, et plus récemment dans les régions étudiées. Cornelis van Lit a ainsi mis en relief le changement de paradigme que connaissent les disciplines du manuscrit dans son *Among Digitized Manuscripts: Philology, Codicology, Paleography in a Digital World* (van Lit, 2020). Dans la lignée de certaines conclusions dressées par l'ouvrage collectif dirigé par E. Muhanna, van Lit montre notamment que les humanités numériques ne s'opposent en rien aux méthodes et aux fins des humanités classiques (philologie, codicologie et paléographie notamment) mais produisent de nouveaux objets (le manuscrit numérisé, par exemple) qui impliquent de nouveaux protocoles de description et d'étude. Si la philologie « digitale » ne s'oppose pas à son aînée classique, les deux doivent prendre acte du bouleversement majeur induit par l'arrivée des techniques numériques. Pour autant, l'apport de celles-ci ne saurait se limiter aux disciplines du manuscrit.

L'intégration d'outils informatiques dans le champ des études historiques a permis aux chercheurs une exploitation plus rapide et systématique des données. Dès les années 1970,

9 <https://www.youtube.com/watch?v=h6UjhqJIsVE>

10 Ce terme fait l'objet de débats, mais fait référence à la lecture de grands corpus grâce à des techniques computationnelles.

l'application des méthodes prosopographiques aux études historiques sur les périodes médiévale et prémoderne s'est montrée particulièrement propice à l'exploitation de la littérature historico-biographique arabe (*ṭabaqāt*), offrant un nouveau regard sur ces sources¹¹. Depuis, de nombreuses bases de données ont vu le jour – d'abord le *Netherlands Ulama Project* sous la direction de Monique Bernards et John Nawas dans les années 1990, suivi par l'*Onomasticum Arabicum* en France¹² ou encore la *Prosopografía de los ulemas de al-Andalus* en Espagne¹³. Les chercheurs n'ont cessé de réfléchir aux enjeux des méthodes statistiques pour écrire l'histoire de l'Islam (voir, entre autres : Bulliet, 1970 ; Romanov, 2017), entre études de cas pratiques et illustrations générales. Ces premières exploitations ont permis d'avancer des hypothèses importantes sur les rythmes et la géographie de l'islamisation dans les premiers siècles de l'islam, ou, par la visualisation des réseaux savants, sur l'évolution de leur organisation à grande échelle.

En effet, la numérisation d'une part et l'émergence d'importantes bases de données d'autre part ont permis aux chercheurs de parcourir de vastes corpus de textes classiques à une vitesse nouvelle et au service de recherches transversales. À mesure que les référentiels textuels devenaient de plus en plus vastes, une « lecture distante » des sources s'est développée¹⁴. Celle-ci induit une rupture incontournable provoquée par les humanités numériques ; les chercheurs, essentiellement historiens, insistent toutefois sur le fait que la technologie ne se substitue pas au travail humain et qu'elle apporte, en plus de nouveaux atouts, son lot de problèmes. La présence de très vastes corpus en ligne ne doit pas donner l'illusion d'une fausse totalité sans interroger ses critères de constitution ; de même, il arrive que les calculs massifs ne fassent que confirmer, au prix de longs efforts, des constats déjà connus de la littérature scientifique. La constitution de corpus contrôlés et la mise en place d'outils performants au service de questions classiques permet, comme pour la littérature juridique en Islam, d'aboutir à des résultats qui confirment parfois des faits établis mais peuvent également en corriger d'autres (Lange et al., 2021).

Si la technologie numérique est souvent mise au service de questionnements classiques (philologiques, historiques, littéraires...), elle crée enfin de nouveaux terrains et de nouveaux types d'enquêtes. Les données produites sur internet, et notamment sur les réseaux sociaux et

11 Voir notamment Bulliet, 1972 ; Chamberlain, 1994. Pour le champ académique francophone, voir Van Rentherghem, 2015.

12 <https://onomasticon.irht.cnrs.fr/>

13 <https://www.eea.csic.es/pua/>

14 Sur le concept de « lecture distante », voir Moretti, 2005 ; Moretti, 2013 ; Jockers, 2013. Pour un exemple de lecture distance appliquée aux mondes musulmans voir Romanov, 2017, où l'auteur explore l'essor et le déclin de tribus et de toponymes sur la base d'une analyse computationnelle de l'un des plus vastes dictionnaires biographiques arabes, le *Ta'rikh al-islām* de Shams al-Dīn al-Dhahabī (m. 748/1348).

les plateformes, ont fait l'objet de nombreux travaux, sans atteindre encore la densité des réflexions méthodologiques proposées au sein des disciplines du manuscrit. La Tunisie et l'Égypte sont parmi les premiers pays être apparus comme des terrains privilégiés des sciences du contemporain (histoire contemporaine, science politique, sociologie) pour faire d'internet et de ses usages un objet d'étude à part entière, en particulier dans l'étude des mouvements révolutionnaires post-2011¹⁵. Plus généralement, le monde arabe est un terrain privilégié pour réfléchir à « l'archivage numérique » et à la conservation des patrimoines en danger face aux conflits, pour étudier à distance des terrains inaccessibles ou pour sauvegarder la mémoire d'événements politiques qui se déroulent en partie en ligne (Bayoumi et Oliveau, 2020).

L'état de la littérature et l'orientation des recherches révèlent enfin une centralité de la langue arabe dans l'analyse des rapports qui se sont noués entre humanités numériques et études des mondes musulmans, pour des raisons qui tiennent en partie à l'investissement massif sur internet, par des acteurs religieux ou non, de ressources numérisées du patrimoine scripturaire islamique, ainsi qu'à l'investissement prononcé des spécialistes de l'Islam classique ou d'islamologie. Cette centralité tient également à la politique active, dans ce domaine, des pays du Golfe, même si la Turquie, par exemple, occupe une place tout aussi centrale ces dernières décennies : la numérisation précoce des archives ottomanes en offrant une illustration, et matière à réflexion pour les spécialistes (Gratien et al., 2014). Si le numérique contribue ainsi à modifier les équilibres linguistiques au sein des mondes musulmans, plusieurs contributions de ce dossier montrent le caractère transversal, dans d'autres espaces, des questions posées par les humanités numériques. Le travail mené par la communauté scientifique française, ces dernières années, illustre également le besoin d'une perspective commune pour tous les spécialistes.

En France, un nouvel élan institutionnel et de nouveaux enjeux pour la recherche collective

Si l'essentiel de l'innovation dans le champ scientifique provient d'abord du monde anglophone, les acteurs de la communauté scientifique française sur les mondes musulmans ont consenti récemment un effort important pour compenser leur retard en la matière. Cet effort se trouvait au croisement de deux priorités structurantes des politiques publiques de la recherche de ces deux dernières décennies : les études aréales et l'accompagnement de la « numérisation du monde » (Ruggiu, 2021). La direction du Groupement d'intérêt scientifique Moyen-Orient et mondes musulmans (GIS-MOMM) a consacré, à partir de 2017, un chantier prioritaire aux humanités numériques, confié à Mercedes Volait, avec l'appui de

15 Pour une vue d'ensemble, mais tout de même centrée sur le Moyen-Orient (Égypte, Jordanie, Koweït), voir Wheeler, 2017. Pour la Tunisie, voir notamment Lecomte, 2013 ou Zayani, 2015. Pour l'Égypte, voir notamment Sakr, 2013 ; Faris, 2013.

postdoctorants, dont Noémie Lucas¹⁶. Sous leur égide, un livre blanc est venu compléter le premier bilan sur les études sur le Moyen-Orient paru en 2014 sous la coordination de Catherine Mayeur-Jaouen (Mayeur-Jaouen, 2014)¹⁷. Il proposait une précieuse photographie de l'état du champ et des projets en humanités numériques, dans un paysage nécessairement mouvant, en se concentrant sur les données textuelles et les disciplines de l'érudition, de fait en avance sur l'usage des outils numériques (Volait et Lucas, 2020). Depuis, une vue d'ensemble de ces mêmes usages pour les sciences du contemporain (sociologie, anthropologie, science politique...) demeure nécessaire.

Le bilan dressé par cette enquête est contrasté : la mise en ligne de ressources numériques est un point fort de la France, grâce aux développements de politiques de numérisation et d'outils de stockage performants – avec la réserve qu'ils sont financés essentiellement par des instruments externes portés par les projets et non sur les ressources propres des institutions, trop insuffisantes, posant ainsi le problème de la pérennité des réalisations numériques. Cet atout est également dû à l'engagement d'un ensemble d'acteurs essentiels du champ : la Maison méditerranéenne des sciences de l'homme (MMSH), l'Institut de recherche sur l'histoire des textes (IRHT), le laboratoire InVisu, ou des bibliothèques comme la Bibliothèque universitaire des langues et civilisations (BULAC) ou la Bibliothèque nationale de France (BnF). À l'étranger, les trois instituts de recherche français de Beyrouth (IFPO), du Caire (CEDEJ) et d'Istanbul (IFEA) ont également conduit des politiques volontaristes. Du côté des difficultés, le travail automatisé sur les textes arabes restait, en 2020, encore balbutiant, résultat aussi de la faiblesse des études littéraires arabes en France. Enfin, l'ensemble des initiatives du champ manquait encore de coordination, et l'activité scientifique de visibilité à l'échelle internationale. Certaines des recommandations de ce rapport ont été suivies de rapides effets : le consortium Huma-Num DISTAM¹⁸ a permis de fédérer une communauté de spécialistes d'études aréales (mondes musulmans, Asie, Afrique) autour des enjeux numériques, des outils de reconnaissance automatique de l'arabe maghrébin ont été développés au sein du GIS-MOMM par deux programmes successifs¹⁹, au sein d'une réflexion sur les manuscrits arabes et le numérique²⁰. Un important volant de formations a

16 Ces deux dernières ont été interviewées par les coordinateurs du numéro, qui les remercient vivement pour cet entretien qui figure à la fin de ce volume.

17 Le rapport est disponible en ligne : <http://majlis-remomm.fr/livre-blanc>

18 DISTAM étant l'acronyme de Digital Studies Africa, Asia, Middle East. Voir le carnet Hypothèses du consortium : <https://distam.hypotheses.org/>

19 Le programme « Khaṭṭ Maġribī », coordonné par Noémie Lucas et Antoine Perrier, a notamment organisé deux hackathons de transcription de manuscrits en arabe maghrébin à la BULAC pour le développement d'un outil HTR de l'arabe, voir l'entretien en fin de dossier.

20 Celle-ci a donné lieu à l'organisation d'une école de printemps sur les humanités numériques et ces mêmes manuscrits, dont les vidéos sont disponibles en ligne. Voir : <https://digimagh.hypotheses.org/194>

permis de familiariser une partie de la communauté scientifique à ces enjeux, faute, pour l'instant, de cursus de licence ou de master spécialisés pour les études aréales²¹. Un rapport final a enfin été publié sur les techniques de reconnaissance automatique des textes manuscrits (HTR) et imprimés (*Optical Character Recognition, OCR*) (Lucas, 2022)²².

Dans ce nouveau cadre, un ensemble de questions traverse les réflexions des acteurs du monde de la recherche – chercheurs, ingénieurs, bibliothécaires – dans les différents espaces qui les réunissent désormais, notamment au sein du consortium DISTAM. Nous voudrions en donner un bref témoignage, pour donner à voir les premiers éléments d'une réflexion collective dans laquelle ce dossier s'inscrit. Devant les moyens considérables qu'offrent désormais les outils numériques, les chercheurs sont séduits par la double tentation de l'exhaustivité et de la vitesse, au risque d'une dispersion des efforts ou de protocoles de recherche mal conçus et, à la fin, contreproductifs ; c'est pourquoi le besoin de coordonner les efforts et les instruments s'impose de manière particulièrement prononcée pour les humanités numériques.

La communauté scientifique française est tout d'abord soumise à un cadre légal qui détermine plusieurs exigences dont le respect est devenu indispensable pour l'allocation de crédits publics dans le cadre de la recherche sur projet. La loi pour une République numérique du 7 octobre 2016, suivie deux ans plus tard d'un Plan national pour la science ouverte, impose aux chercheurs l'ouverture et le libre partage de leurs données et des publications scientifiques. Cette obligation vise non seulement l'introduction de transparence dans un monde informatique fortement dominé par des logiques économiques ou corporatistes, à travers la généralisation d'un format de données « FAIR²³ », mais elle sert également à renforcer les liens entre société et communauté scientifique, en rendant accessible à tous le travail de la recherche publique. Une telle logique est également destinée à renforcer la visibilité internationale des travaux français et à combattre les logiques de privatisation, particulièrement vigoureuses dans le contexte anglophone, qui restreignent l'accès aux publications derrière des portails onéreux. Pour les études sur les mondes musulmans, la garantie d'une ouverture est d'autant plus précieuse à l'heure où de nombreux régimes de la région font peser sur le web surveillance ou censure. Elle l'est aussi dans le contexte de

21 Quelques exemples : les journées de formation sur l'édition et la publication de textes arabes avec le numérique en décembre 2020 (l'ensemble des enregistrements est disponible sur la page YouTube du GIS-MOMM : <https://www.youtube.com/@gismomm777/videos>), d'autres journées sur les études sur le Maghreb et le numérique en mars 2021 (voir également la page YouTube du GIS-MOMM) ; ou l'Action nationale de formation du CNRS DIGITAL AREA en novembre 2021 (voir : <https://www.gis-reseau-asie.org/evenement/une-action-nationale-de-formation-en-humanites-numeriques-pour-les-etudes-areales>).

22 Pour le carnet de recherche sur la philologie numérique, voir : <https://philaranum.hypotheses.org/>. Le rapport sur les techniques OCR et HTR de l'arabe manuscrit est disponible en ligne : <http://majlis-remomm.fr/72481>

23 Faciles à trouver, accessibles, interopérables et réutilisables.

« langues peu dotées » dans ces espaces, c'est-à-dire disposant d'une quantité plus faible de données en ligne. Leur présence massive est pourtant indispensable pour l'amélioration des outils qui en facilitent le traitement, de la traduction à la lexicométrie, et même pour l'usage d'instruments qui seront ou sont déjà systématiquement associés à la recherche financée. Les chercheurs doivent dès lors se familiariser avec de nouveaux types de publication pour partager leurs données, comme le *datapaper* qui présente ces données sous un format éditorialisé, ou prévoir, dans chaque projet de recherche, un plan de gestion de données.

L'ouverture des données rejoint directement l'enjeu, central, de la façon dont elles sont conservées, stockées, affichées et répertoriées sur le web. Si les infrastructures publiques comme Huma-Num proposent un ensemble de solutions aux chercheurs pour déposer leurs données de manière sécurisée et les enrichir pour en faciliter la visibilité²⁴, la question de la pérennité de celles-ci se pose de manière instante dans un environnement numérique où la durée de vie des pages internet est brève. Ce besoin dépasse le simple enjeu du stockage, pose la question des infrastructures publiques disponibles pour en garantir la sécurité²⁵ mais aussi celle des modes de financement de la recherche. Les financements sur projet, qui permettent en général de produire d'importantes quantités de données, ne prévoient pas toujours de relais, en ressources humaines comme en infrastructure numérique, pour garantir l'accessibilité à long terme des données, qui risquent ainsi de disparaître ou d'être englouties dans les profondeurs du web.

Ces enjeux n'ont rien de spécifique aux études sur les mondes musulmans et concernent toutes les sciences humaines. Certaines difficultés sont néanmoins plus prononcées ; ainsi de l'établissement d'un langage commun au sein des projets. L'emploi de langues écrites dans un alphabet non latin, la variation des translittérations ou de l'onomastique posent des difficultés supplémentaires au besoin impérieux d'alignement des référentiels, autrement dit de classification des données au moyen d'un même vocabulaire, pour en garantir l'accès et prévenir la multiplication des doublons²⁶. La description numérique des archives, images, publications numériques est comme un langage supplémentaire qui s'ajoute aux nombreuses langues déjà pratiquées dans les mondes musulmans ; ce langage réclame également un grand effort d'uniformisation, à l'échelle nationale mais plus généralement au sein de la communauté scientifique mondiale. Ce besoin rejoint celui d'une meilleure coordination entre des métiers trop séparés par les institutions françaises : chercheurs, bibliothécaires ou ingénieurs, dont la répartition des tâches est repensée par les humanités numériques. Les

24 <https://www.huma-num.fr/>

25 Le Centre informatique national de l'enseignement supérieur (CINES), qui offre ces garanties, ne rend pas les données visibles et accessibles.

26 La réflexion sur ces alignements est au cœur du projet Mistara (« Onomastique arabe et métadonnées des langues à écriture arabe dans les référentiels en ligne »), qui a tenu un carnet Hypothèses : <https://mistara.hypotheses.org/>.

nombreux projets qui se sont développés dans le champ ont montré que les outils numériques ne sont plus cet instrument subsidiaire appelé à répondre à des questions scientifiques élaborées de manière préalable et indépendante ; sans constituer une fin en soi, ils doivent être intégrés dès le départ à la confection d'un projet et à l'élaboration de problématiques de manière concertée entre métiers. Intégrer les potentialités numériques dans les tous premiers moments de la recherche, notamment doctorale, réclame encore la généralisation de l'enseignement en humanités numériques dans les cursus de licence, y compris dans les départements de langues ou les cursus spécialisés, par exemple en islamologie. Celui-ci demeure rare, à de remarquables exceptions près, comme l'École des chartes²⁷.

En France, les portails de ressources numériques et les outils de signalement des ressources se multiplient, y compris dans le domaine des mondes musulmans, grâce à des portails aréaux comme celui sur le Maghreb²⁸, des blogs, journaux ou carnets de recherche des projets financés, qui deviennent eux-mêmes des résultats de recherche²⁹ ou des portails de ressources en ligne³⁰. Les archives françaises et les bibliothèques ont également pris le tournant de la numérisation progressive de leurs ressources, consultables en libre accès (*Open Access*), et de la mise en ligne d'instruments de recherche³¹. En cela, les pays du Maghreb et du Moyen-Orient suivent un chemin similaire, des politiques patrimoniales particulièrement dynamiques conduisant à la numérisation de plus en plus massive de textes et, plus rarement, d'images, sans qu'elles ne s'accompagnent nécessairement de leur mise en ligne en accès libre. La rencontre consacrée aux humanités numériques au sein du Forum Insaniyyat Tunis 2022 (Bendana et al., 2023) témoignait d'une certaine vitalité de ces préoccupations en Afrique du Nord, du Maroc à l'Égypte, avec deux impulsions essentielles données par les institutions patrimoniales, en particulier les bibliothèques et les universités, notamment tunisiennes, où la formation au numérique se développe progressivement³². Dans le même temps, le web est

27 Seule université française qui, à ce jour, propose un master « Humanités numériques et sciences sociales ».

28 <https://etudesmaghrebines.cnrs.fr/>

29 On peut prendre l'exemple des billets du « Lexique vivant de la révolution et de la guerre en Syrie » de l'ANR SHAKK (Syrie : conflits, déplacements, incertitudes) : <https://syria-lexicon.pubpub.org/>.

30 À titre d'exemples, la BINA, bibliothèque numérique de la BULAC, héberge des manuscrits numérisés dont certains sont regroupés en parcours, comme ceux du projet TariMa sur le Maghreb (<https://bina.bulac.fr/s/tarima/page/welcome>), la Perséide Athar (<https://athar.persee.fr/>), ou le portail « Bibliothèques d'Orient » de la Bibliothèque nationale (<https://www.bnf.fr/fr/bibliotheques-dorient>). On trouve aussi des ressources numériques sur le Maghreb et le Moyen-Orient sur des plateformes au périmètre plus vaste comme Biblissima (<https://portail.biblissima.fr/en>).

31 Effort consenti par les Archives nationales de la France d'outre-mer ou les Archives diplomatiques, qui mettent en ligne des instruments de recherche et des ressources numérisées.

32 En témoigne le projet Raqmyat, mentionné dans l'entretien à la fin de ce dossier.

regardé avec méfiance par des régimes autoritaires qui surveillent ce lieu d'expression privilégié de la contestation. Des bilans manquent encore pour établir la percée et les particularités de la « transition numérique » dans les mondes musulmans et pour mesurer ses conséquences sur un possible rapprochement entre les pays du Nord et ceux du Sud. Si certaines plateformes de mise à disposition de ressources dans les langues de la région, principalement en arabe, sur des sites eux-mêmes entièrement en arabe, comme la bibliothèque de textes al-Maktaba al-Shāmila³³, réunissent des chercheurs du monde entier, la fracture numérique et la disparité des financements font aussi courir le risque d'écart croissants entre les communautés scientifiques.

Le livre blanc de 2020 soulignait le fait qu'en dépit d'un investissement réel de la communauté scientifique française, « le numérique est cependant surtout envisagé comme un médium plutôt que comme un vecteur de transformation méthodologique profonde » (Volait et Lucas, 2020 : 94). Les spécialistes des mondes musulmans ont parfois tendance à accueillir de manière prudente des changements numériques qui rappellent d'anciennes menaces : la mise en ligne de ressources en langues vernaculaires ressuscite la tentation d'un orientalisme distant, qui sépare lecture des textes et fréquentation du terrain (Jomier et Warscheid, 2018 : 314), et concentre la première sur une sélection arbitraire de manuscrits disponibles en ligne, tant les propriétaires de textes plus rares numérisent volontiers leurs documents pour les conserver, sans pour autant les rendre accessibles en ligne³⁴. Les disciplines de l'érudition, tout en se distinguant dans l'usage des humanités numériques, abritent également de nombreuses résistances, défendant la conviction que la machine ne se substituera pas à la main humaine pour éditer et lire des textes, et qu'un survol de gigantesques corpus numérisés ne se substituera pas à la fréquentation patiente d'un texte tout entier. Si aucun spécialiste n'est assez naïf pour croire qu'une méthode se substituera tout à fait à l'autre, la séduction de lectures plus rapides, aux résultats d'apparence plus spectaculaire, est un danger. Il manque encore au tournant impliqué par les humanités numériques une réflexion sur le réagencement global entre savoir, société et technique, dans une perspective proche de celle de Bruno Latour, réflexion qui serait peut-être de nature à prévenir les inquiétudes. De la même façon que l'imprimerie n'a pas toujours représenté une rupture immédiate dans la nature des œuvres produites dans le monde arabe (par exemple au Maroc : Abdulrazak, 1990), rien ne prédit encore que la numérisation massive des ressources textuelles n'en modifiera la production ou n'en favorisera la circulation dans le monde. De la même façon, l'archivage numérique n'est pas toujours un moyen de réduire les inégalités entre Nord et Sud mais peut au contraire les aggraver (Husaini et Maingraud-Martinaud, 2024). La nature ambivalente de ces changements n'en diminue pas l'importance, mais celle-ci doit encore être pensée. En apportant un très

33 <https://shamela.ws/>

34 Voir, par exemple, à ce sujet : Voguet, 2018.

modeste apport à ce débat, les contributions de ce dossier montrent la fécondité d'approches à la fois engagées et critiques dans l'usage des humanités numériques.

Les contributions du dossier, exemples d'usages scientifiques des humanités numériques

Les contributions de ce dossier relèvent, comme le bilan mené par Mercedes Volait et Noémie Lucas, essentiellement des disciplines proches des textes ou de spécialistes réfléchissant du moins à des questions d'archivage. Elles reflètent l'état des propositions reçues et d'une avance, certainement provisoire, de cette partie des disciplines par rapport à la sociologie ou à la science politique, peu représentées ici, à l'exception de l'article de Véronique Ginouvès et Ibrahim Moussa. La contribution de Sophia Mahroug, historienne, montre toutefois le caractère interdisciplinaire des réflexions sur le numérique. L'ensemble du dossier met également en relief l'importance de la recherche collective pour les humanités numériques, puisqu'elles s'adossent souvent à des projets financés par l'Agence nationale de la recherche (ANR), le Conseil européen de la recherche (European Research Council, ERC) ou le GIS Collex-Persée. La diversité des métiers est également frappante et se retrouve dans ces textes, écrits pour la plupart à plusieurs mains : calligraphes professionnels et historiens de l'art, sociologues et archivistes, historiens et archivistes ou bibliothécaires. Ils représentent assez fidèlement la diversité des corpus concernés par les humanités numériques : images (V. Thérouin et al.), textes (A. Perrier), sons (V. Ginouvès et I. Moussa), corpus nativement numériques ou disponibles désormais seulement sur internet, pour le cas iranien (N. Pashkeeva et al. ; S. Mahroug). Tout en étant limités à des espaces essentiellement maghrébins et moyen-orientaux, sans représentation de disciplines importantes en ces matières, comme la littérature arabe, ils offrent un premier paysage sinon représentatif du moins illustratif, espérons-le, des usages scientifiques des humanités numériques au sein de ce champ.

La demande adressée aux auteurs consistait à exposer le protocole critique de leur usage des outils numériques. Celui-ci commençait naturellement par des problèmes d'ordre technique ou, autrement dit, par des questions posées par les chercheurs aux technologies. La façon d'organiser et surtout d'enrichir les données par un vocabulaire adéquat – instrument et résultat même de la recherche – traverse plusieurs articles : trouver la bonne catégorie de descripteurs des calligraphies arabes (V. Thérouin et al.) ou des archives disponibles sur le web (N. Pashkeeva et al.), modéliser une base de données pour la rendre à la fois interopérable et adaptée aux besoins spécifiques d'un projet de recherche, délimiter le vocabulaire le plus pertinent pour la fouille de texte (A. Perrier). Les auteurs doivent alors s'adapter à des langages qui ont tendance à restreindre la diversité d'une société ou de ses représentations à une liste fermée de mots. Le format même des archives est interrogé par les articles, qui évoquent le passage successif du mode analogique au microfilm puis à la numérisation (M. Jebahi ; V. Ginouvès et I. Moussa), passage qui doit être accompagné d'une

mise en contexte indispensable pour l'intelligence d'un fonds – un souci constant des archivistes. La question du format est d'autant plus instante à l'heure de sa fragilité, quand une page web peut disparaître à tout moment et doit être archivée sous un format adéquat (S. Mahroug). Des exemples d'usage d'outils sont également évoqués, à travers le recours au langage Dublin Core, aux technologies HTR de Calfa ou aux plateformes comme Omeka, sans s'attarder sur leurs dimensions techniques.

Les contributions approfondissent en revanche les résultats scientifiques qui se divisent, schématiquement, entre des découvertes issues de la fouille ou de l'organisation nouvelle de données déjà existantes et l'apport de nouvelles sources par l'intermédiaire du numérique. Dans la première catégorie, la faculté nouvelle de traiter de grands corpus de données, étalés dans l'espace, le temps ou entre les genres, permet d'explorer des calligraphies éloignées d'une centralité canonique au profit du genre des « frontières » (V. Théroutin et al.). Différences et langages communs peuvent être identifiés grâce à la fouille de texte dans des sources issues de régions ou de genres différents, au sein du Maroc moderne par exemple (A. Perrier). L'extraction automatique de texte à partir de manuscrits conservés en France facilite l'accès à des textes souvent déjà édités localement mais indisponibles au format numérique. Ce format permet aussi la redécouverte de documents, comme les archives sonores, dont la numérisation permet de « faire entendre de nouveau [d]es voix », celles des individus, objets d'enquêtes anthropologiques au Niger, ou des anthropologues eux-mêmes (V. Ginouvès et I. Moussa). La présence de données nativement numériques ou bien présentes exclusivement sur internet, faute d'autres archives institutionnelles, permet de reconstituer la sociologie d'un corps armé inaccessible, comme celui des Gardiens de la révolution (S. Mahroug) ou bien de documenter l'histoire de la violence révolutionnaire en Iran en contrant les récits officiels (N. Pashkeeva et al.). Dans le cas iranien, internet est le lieu de stockage d'archives qui lui préexistent, mais il offre également un lieu alternatif de conservation aux institutions d'État.

Dans tous les cas, les auteurs des articles ne se laissent pas envoûter par la fausse clarté d'archives disponibles mais décontextualisées, parfois manipulées ou problématiques (N. Pashkeeva et al. ; S. Mahroug), ou sont conscients des limites d'une lecture distante de sources, au détriment d'une lecture exhaustive de celles-ci (A. Perrier). Les limites, soulignées, ne découragent pas pour autant l'emploi de nouvelles techniques, associées à d'anciens modes de lecture. Du reste, ceux-ci se retrouvent dans l'analyse de sources par l'historien, indépendamment de leur format, ou dans une démarche ethnographique (N. Pashkeeva et al.), et la recherche d'« indices » (S. Mahroug) trouve une égale satisfaction sur internet que dans des archives plus classiques. Dans le même esprit, d'autres contributions inscrivent le saut numérique dans une histoire plus longue de l'archivage et des techniques de conservation, faisant remonter la rupture, plus décisive peut-être que la « numérisation du monde » de ces deux dernières décennies, à la généralisation de l'informatique dans les années 1980 (M. Jebahi ; V. Ginouvès et I. Moussa).

Les articles de ce numéro mettent enfin en relief quelques particularités des études aréales. Les modes de traitement de l'écriture arabe précèdent la fouille de texte ou la description des calligraphies, bien que, comme le souligne Noémie Lucas, les recherches s'inscrivent dans un contexte où l'arabe fait de moins en moins figure de langue « peu dotée ». Les particularités soulevées dans ces articles s'inscrivent plus profondément dans le contexte même des sociétés africaines, maghrébines ou moyen-orientales étudiées. La séquence post-2011, ou plus spécifiquement post-1979 pour l'Iran, traverse les contributions, sans que celles-ci ne renouent avec une sorte de « déterminisme numérique » destiné, dans un enthousiasme initial, à expliquer les « printemps arabes » par internet et les réseaux sociaux (S. Mahroug). Le développement d'une justice transitionnelle en Tunisie, de la même façon qu'au Maroc, a donné un nouvel élan à l'archivage et à la numérisation (M. Jebahi), où le numérique répond aussi, à sa place, à un « besoin d'histoire ». Dans le cas de terrains fermés par la guerre ou les régimes autoritaires, le numérique est devenu indispensable pour se substituer à la fréquentation empêchée du terrain.

Les ressources en ligne sont restituées dans leur ambivalence s'agissant d'un rapprochement entre pays du Nord et pays du Sud, rendu compliqué par le contexte de fortes tensions dans une partie de ces régions. Le web est un espace où s'expriment simultanément les voix des opposants à des régimes résolus à leur en barrer l'accès et les récits manipulés par les États. Il est aussi une ressource bien rare dans des pays pour certains majoritairement ruraux, où l'accès à internet est inégal, comme le Niger (V. Ginouvès et I. Moussa). Les appareils critiques ou les langages de description accordent une place de choix aux langues européennes, singulièrement l'anglais, rarement lues par une part significative des populations de la région. Toutefois, ces contributions montrent également des espaces de dialogue, à travers des formes de partenariat et de coopération autour des techniques informatiques en Tunisie (M. Jebahi) ou dans l'usage, par les chercheurs français, de ressources localisées au Maghreb, comme celles de la fondation du Roi Abdulaziz al Saoud, riches et facilement accessibles en ligne (A. Perrier). Face à des sujets sensibles, comme la restitution des archives aux pays anciennement colonisés, la numérisation permet de transformer, par de nouveaux usages, un matériel produit en contexte colonial (V. Ginouvès et I. Moussa), de rapatrier au moins une copie des archives (M. Jebahi), ou d'imaginer une étape où pourrait être dépassée la question de la présence physique des documents au profit d'un patrimoine numériquement accessible désormais en partage. En offrant quelques pistes de réflexion propres aux études sur les mondes musulmans, ce bilan critique de coordinateurs de projets de recherche sur les humanités numériques est, enfin, un moyen de réfléchir aux transformations actuelles et rapides des métiers des sciences humaines et sociales en général.

Bibliographie

ABDULRAZAK Faouzi, 1990, *The Kingdom of the Book: The History of Printing as an Agency of Change in Morocco between 1865 and 1912*, thèse de doctorat, Boston University.

- BAYOUMI Hala et OLIVEAU Sébastien, 2020, « Digital archiving in the Arab world: Assessment and challenges », *Égypte/Monde arabe*, 22, p. 13-20.
- BENDANA Kmar, HMED Choukri, PERRIER Antoine et VALLET Éric, 2023, « “Tunis, capitale des sciences humaines et sociales”. Retour sur le Forum Insaniyyat Tunis 2022 : les conditions d’un partenariat scientifique Nord-Sud », *L’Année du Maghreb*, 30.
- BULLIET Richard, 1972, *The Patricians of Nishapur. A study in medieval Islamic social history*, Cambridge, Harvard University Press.
- BULLIET Richard W., « A Quantitative Approach to Medieval Muslim Biographical Dictionaries », *Journal of the Economic and Social History of the Orient*, 13(2), p. 195-211.
- CHAMBERLAIN Michael, 1994, *Knowledge and social practice in medieval Damascus, 1190-1350*, Cambridge, Cambridge University Press.
- FARIS David, *Dissent and Revolution in a Digital Age: Social Media, Blogging and Activism in Egypt*, New-York, I.B. Tauris.
- GOLD Matthew K., 2012, *Debates in the Digital Humanities*, Minneapolis, University of Minnesota Press.
- GRATIEN Chris, POLCZYŃSKI Michael et SHAFIR Nir, 2014, « Digital Frontiers of Ottoman Studies », *Journal of the Ottoman and Turkish Studies Association*, 1(1-2), p. 37-51.
- HUSAINI Sa’eed et MAINGRAUD-MARTINAUD Cyrielle, 2024, « Asymétries transnationales dans l’archivage de l’activisme », *Critique internationale*, 102(1), p. 121-143.
- JOCKERS Matt, 2013, *Macro-analysis: Digital methods and literary History*, Urbana, University of Illinois Press.
- JOMIER Augustin et WARSCHEID Ismail, 2018 « Pour une islamologie historique », *Annales. Histoire, Sciences sociales*, 73, p. 311-316.
- LANGE Christian et al., 2021, « Text Mining Islamic Law », *Islamic Law and Society*, 28, p. 234-281.
- LECOMTE Romain, 2013, « Expression politique et activisme en ligne en contexte autoritaire. Une analyse du cas tunisien », *Réseaux*, 181, p. 51-86.
- LUCAS Noémie, 2022, « OCR/HTR et graphie arabe. Les manuscrits arabes à l’heure de la reconnaissance automatique des écritures », *Cahier du GIS Moyen-Orient et mondes musulmans*, 3.

MAYEUR-JAOUEN Catherine, 2014, *Livre blanc des études françaises sur le Moyen-Orient et les mondes musulmans*, livre blanc, GIS Moyen-Orient et mondes musulmans, 2020.

MORETTI Franco, 2005, *Graphs, Maps, Trees: Abstract Models for a Literary History*, Londres, Verso.

MORETTI Franco, 2013, *Distant Reading*, Londres, Verso.

MUHANNA Elias (dir.), 2017, *The Digital Humanities and Islamic & Middle East Studies*, Berlin-Boston, De Gruyter.

ROMANOV Maxim, 2017, « Algorithmic Analysis of Medieval Arabic Biographical Collections », *Speculum*, 92, p. 226-246.

ROMANOV Maxim, 2017, « Toward Abstract Models for Islamic History », in MUHANNA Elias (dir.), *The Digital Humanities and Islamic & Middle East Studies*, Berlin-Boston, De Gruyter.

RUGGIU François-Joseph (dir.), 2021 *Les études aréales*, Paris, Athéna.

SAKR Laila Shereen, 2013, « A Digital Humanities Approach: Text, the Internet, and the Egyptian Uprising », *Middle East Critique*, 22(3), p. 247-263.

VAN LIT Cornelis, 2020, *Among Digitized Manuscripts: Philology, Codicology, Paleography in a Digital World*, Leyde-Boston, Brill.

VAN LIT Cornelis, MORRIS James H. et CEVIK Deniz, 2020 « A Digital Revival of Oriental Studies », *Revue de l'Institut des langues et cultures d'Europe, Amérique, Afrique, Asie et Australie*, 39, [en ligne] <https://doi.org/10.4000/ilcea.10327>.

VAN RENTHERGHEM Vanessa, 2015, *Les élites bagdadiennes au temps des Seldjoukides. Étude d'histoire sociale*, Damas-Beyrouth, Institut français du Proche-Orient.

VOGUET Élise, 2018, « Travailler sur les manuscrits du Touat : expérience de recherche dans les bibliothèques privées (*khizānāt*) du Sahara algérien », in LAROQUE Claude (dir.), *La peinture et l'écrit au Moyen-Orient, supports et tracés*, Paris, HiCSA.

VOLAIT Mercedes et LUCAS Noémie, 2020, *Vers la science ouverte ? La transition numérique et la recherche sur le Moyen-Orient et les mondes musulmans en France*, livre blanc, GIS Moyen-Orient et mondes musulmans, 2020.

WERTHNER Hannes et al., 2024, *Introduction to Digital Humanism. A Textbook*, Cham, Springer.

WHEELER Deborah L., 2017, *Digital Resistance in the Middle East: New Media Activism in Everyday Life*, Édimbourg, Edinburgh University Press.

ZAYANI Mohammed, *Networked publics and digital contention. The politics of every life in Tunisia*, New-York, Oxford University Press.