

APPORT DU LOGICIEL LIBRE EN IMAGERIE MÉDICALE

JODOGNE S (1)

RÉSUMÉ : La recherche médicale recourt à des données de plus en plus massives, complexes et interdépendantes dont l'analyse nécessite l'usage d'algorithmes spécialisés. Afin de reproduire et valider les résultats d'une étude scientifique de manière indépendante, il n'est, dès lors, plus suffisant de partager le texte de l'article en «open-access» complété avec les données brutes en «open-data». Il convient désormais d'également partager les algorithmes qui ont servi à l'analyse des données avec d'autres équipes de chercheurs. Le logiciel libre et «open-source» répond précisément à ce besoin de diffuser les connaissances techniques à grande échelle. Dans cet article, nous présentons plusieurs exemples de logiciels libres utilisés en médecine, avec une attention particulière portée à l'imagerie médicale.

MOTS-CLÉS : *Reproductibilité - Logiciel libre - Open-source - Imagerie médicale*

CONTRIBUTION OF FREE SOFTWARE TO MEDICAL IMAGING

SUMMARY : Medical research uses increasingly massive, complex and interdependent data, the analysis of which requires the use of specialized algorithms. In order to independently reproduce and validate the results of a scientific study, it is no longer sufficient to share the text of the article as an open-access document, together with the raw research data according to the open-data approach. It is now also needed to share the algorithms used to analyze the data with other research teams. Free and open-source software precisely responds to this need to disseminate technical knowledge at a large scale. In this paper, we present several examples of free software used in medicine, with a particular focus on medical imaging.

KEYWORDS : *Reproducibility - Free software - Open source - Medical Imaging*

INTRODUCTION

L'open-access est un modèle de diffusion de la recherche scientifique qui vise à rendre les résultats de la recherche librement accessibles en ligne. L'open-access permet notamment à quiconque, qu'il s'agisse d'un chercheur, d'un étudiant, d'un professionnel ou du grand public, d'accéder gratuitement aux publications scientifiques. Cela élimine les barrières financières et garantit que la connaissance scientifique soit accessible à tous, quelle que soit leur situation géographique ou leur affiliation institutionnelle. En rendant la recherche plus largement accessible, l'open-access favorise, en outre, la dissémination rapide des connaissances, la transparence et l'éthique. L'open-access s'est grandement popularisé au cours de la dernière décennie. Des revues en open-access de haute qualité ont été créées, et de nombreuses revues traditionnelles ont commencé à offrir des options d'open-access.

L'open-access est un pas important vers une diffusion plus large et équitable de la recherche, mais il ne résout pas tous les problèmes liés à la production scientifique. En effet, le fait de disposer du texte d'un article scientifique ne permet pas, à lui seul, d'être en mesure de vérifier et de reproduire indépendamment les résultats décrits dans l'article. Or, la reproductibilité est

essentielle dans une démarche scientifique car elle permet de vérifier la validité des résultats expérimentaux, renforçant ainsi la confiance dans la fiabilité des découvertes. La reproductibilité favorise également la collaboration entre chercheurs en permettant à d'autres équipes de s'appuyer sur des travaux antérieurs en toute confiance.

Afin de garantir cette reproductibilité, il convient de compléter l'open-access avec un accès aux données de recherche sous-jacentes. La mise à disposition des données brutes est généralement qualifiée de démarche d'open-data. La gestion et le partage des données de recherche en open-data restent, bien entendu, un défi, singulièrement dans le domaine médical, en raison de la sensibilité des données personnelles des patients, ce qui nécessite la mise en place de protocoles de confidentialité et de consentement rigoureux. De plus, la diversité des formats de données et des pratiques de collecte complexifient la standardisation et la mise à disposition de ces informations de manière cohérente à d'autres équipes.

Néanmoins, d'importantes initiatives d'open-data existent dans la recherche médicale. Par exemple, «*The Cancer Imaging Archive*» (TCIA) est une archive en ligne de données d'imagerie médicale liées au cancer, offrant un accès libre et gratuit à des ensembles de données d'imagerie pour la recherche en oncologie (1). Un autre exemple est «*The Cancer Genome Atlas program*» (TCGA), qui a rassemblé des chercheurs et des institutions du monde entier dans le but de séquencer et de caractériser, de manière approfondie, des milliers de tumeurs cancéreuses provenant de patients, et qui a

(1) ICTEAM (Institute of Information and Communication Technologies, Electronics and Applied Mathematics), UCLouvain, Louvain-la-Neuve, Belgique.

mené à la publication de données de type génomique, épigénomique, transcriptomique et protéomique en open-data (2). Sur le plan de la standardisation des données cliniques, il faut souligner l'existence de l'initiative de recherche collaborative OMOP («Observational Medical Outcomes Partnership») qui développe un modèle de données commun pour faciliter la recherche médicale basée sur des données de santé réelles. Le standard OMOP est de plus en plus largement utilisé en support à la recherche clinique dans les hôpitaux européens.

La recherche biologique et clinique s'appuie traditionnellement sur l'analyse qualitative ou quantitative de paramètres mesurables bien identifiés, à travers des méthodes statistiques établies. Comme ces méthodes statistiques sont acceptées par l'entière communauté scientifique, un accès en open-data aux données cliniques permet de garantir la reproductibilité des résultats. Néanmoins, l'évolution des technologies biomédicales et médicales conduit à une production de données de plus en plus massives, complexes et interdépendantes (images médicales, séquençages multi-omiques, textes médicaux, séries temporelles...) qui nécessitent le recours à des algorithmes spécialisés pour pouvoir les croiser et en extraire des paramètres mesurables en amont d'une analyse statistique.

De nos jours, ces algorithmes sont généralement acquis par les chercheurs auprès d'entreprises, ce qui limite la reproductibilité des résultats aux équipes qui disposent du même logiciel d'analyse. En effet, contrairement aux méthodes statistiques, il n'y a généralement pas de consensus sur l'algorithme qu'il convient d'utiliser, et chaque fournisseur offre une solution propriétaire différente. Les détails sur le fonctionnement interne de ces algorithmes propriétaires sont souvent gardés secrets par les entreprises qui les développent. L'émergence de l'intelligence artificielle (IA) en médecine complexifie encore cette situation, en ajoutant une variance très forte liée aux données et à l'architecture du modèle qui ont été utilisées lors de l'entraînement. L'acquisition d'un algorithme donné peut également représenter un obstacle financier important pour de nombreuses institutions, notamment dans les économies émergentes, ce qui peut casser la dynamique de recherche collaborative. Par ailleurs, lorsqu'une institution de recherche clinique ou un établissement de santé dépend d'algorithmes propriétaires, ils peuvent devenir vulnérables aux modifications de tarification, aux changements des conditions de licence, ou même, à la cessation du support technico-commercial, ce qui entraîne une dépendance forte du processus de

recherche vis-à-vis de décisions du fournisseur, ainsi qu'un manque de durabilité. On observe également fréquemment des algorithmes propriétaires qui sont conçus pour ne fonctionner qu'avec des systèmes spécifiques ou des équipements médicaux commercialisés par la même firme. Il existe, enfin, un mouvement global de l'industrie pour déplacer les algorithmes dans le «cloud» afin de pouvoir facturer à l'analyse, ce qui peut soulever des inquiétudes concernant la confidentialité et la sécurisation des données qui sont traitées en-dehors des infrastructures de l'institution de recherche.

On le voit, du fait de la dépendance grandissante de la recherche vis-à-vis des algorithmes, la combinaison de l'open-access avec l'open-data n'est plus suffisante pour garantir une reproductibilité des résultats de recherche. Il convient, désormais, de partager aussi le code source des logiciels qui ont servi à analyser les données brutes avant de produire l'article final. En toute logique, cette démarche de publication du code source est fréquemment qualifiée d'open-source. Lorsque des chercheurs partagent le code source de leurs analyses, cela permet à d'autres équipes de comprendre, en détail, comment les résultats ont été obtenus, de réutiliser le code pour effectuer des analyses similaires et de détecter d'éventuelles erreurs ou biais. La démarche d'open-source est devenue essentielle pour la reproductibilité, car elle garantit que les méthodes et les analyses puissent être examinées, comprises et reproduites par d'autres chercheurs.

Dans cet article, après une définition plus précise du logiciel libre, nous présentons plusieurs exemples de logiciels libres utilisés en médecine, avec une attention particulière portée à l'imagerie médicale. Ce texte ne se veut en aucune manière exhaustif, car l'écosystème de ces logiciels est très riche et en constante évolution.

LOGICIEL LIBRE ET OPEN-SOURCE

Historiquement, jusque dans les années 1970, la vente de matériel informatique était couplée à la vente des logiciels qui tournaient sur ces machines. En effet, dans les premiers jours de l'informatique, le matériel et le logiciel étaient souvent conçus de manière intégrée, car l'architecture matérielle imposait des contraintes spécifiques sur le logiciel. Par la suite, la standardisation des composants matériels et la démocratisation de leur production ont ouvert la voie à un modèle où le matériel était vendu

séparément du logiciel. Cette évolution technologique a donné naissance à l'industrie du logiciel telle que nous la connaissons aujourd'hui, basée sur des modèles économiques comme la vente de licences, la vente de services associés ou, plus récemment, les formules d'abonnement et de tarification à l'utilisation.

L'exploitation commerciale de logiciels est souvent associée à l'idée de restreindre les libertés des utilisateurs de manière à créer une dépendance vis-à-vis des fournisseurs et ainsi maximiser les profits de ces derniers. On parle alors de logiciels propriétaires, voire de logiciels privés. Les restrictions pratiquées se matérialisent au niveau des licences logicielles. Celles-ci peuvent typiquement interdire l'utilisation du logiciel à des fins commerciales, limiter le nombre d'utilisateurs autorisés, ou proscrire la redistribution du logiciel à d'autres personnes ou organisations. En règle générale, dans le modèle propriétaire, le code source du logiciel est gardé secret, ce qui signifie que les utilisateurs n'ont pas accès au code informatique sous-jacent. Cela limite leur capacité à comprendre comment fonctionne le logiciel ou à identifier des problèmes de sécurité potentiels. Dans de nombreux cas, les utilisateurs ne sont pas autorisés à apporter eux-mêmes des changements au programme pour le personnaliser ou l'adapter à leurs besoins spécifiques. Ces restrictions associées au logiciel propriétaire peuvent être problématiques, notamment dans un contexte scientifique.

En réaction aux restrictions imposées par les logiciels propriétaires et à la perte de libertés qu'elles entraînent pour les utilisateurs, l'idée de logiciel libre a été formalisée dans un cadre académique dès le début des années 1980 par Richard Stallman (3). Un logiciel est dit «libre» si, et seulement si, sa licence associée garantit les quatre libertés fondamentales d'utiliser le logiciel pour tout usage, d'étudier et adapter le fonctionnement du logiciel, d'en distribuer des copies, et de distribuer des versions améliorées du logiciel. Il est important de constater que les deuxième et quatrième libertés de cette liste imposent un accès complet au code source du logiciel, d'où le vocable «logiciel libre et open-source». Cette formalisation a donné naissance au «GNU Project», dès 1983, dont le but est de développer un système d'exploitation complet composé exclusivement de logiciels libres, puis à la «Free Software Foundation» (FSF), en 1985, pour soutenir le mouvement global du logiciel libre.

De nombreuses licences logicielles compatibles avec la définition des logiciels libres ont ensuite été créées. On peut distinguer deux

grandes catégories de telles licences : celles dites «copyleft» et celles dites «permissives». Les premières imposent une forme de réciprocité, à savoir que les versions dérivées d'un logiciel libre doivent être publiées sous une licence au moins aussi forte que le logiciel original. Pour cette raison, les licences copyleft interdisent certaines formes de réutilisation du code source dans des logiciels propriétaires. L'objectif est ici d'instaurer un cercle vertueux qui permet de faire grandir un logiciel libre grâce aux contributions d'une communauté de développeurs, tout en préservant la liberté du code source initial. L'exemple prééminent de licence copyleft est la «GNU General Public License» (GPL), qui a été proposée par Richard Stallman dès 1989 en support au GNU Project et qui est, notamment, utilisée par le noyau Linux. Quant aux licences permissives, elles permettent une réutilisation plus flexible du code source, y compris dans des logiciels propriétaires, c'est-à-dire sans l'obligation de rendre le code source dérivé ouvert. Plutôt qu'une réciprocité, les licences permissives visent une dissémination maximale des logiciels libres, y compris au sein de l'industrie. Des exemples de licences permissives incluent les licences Apache, MIT et BSD.

Le choix d'une licence plutôt qu'une autre reflète la stratégie des auteurs d'un logiciel libre. Dans les prochaines sections, différents logiciels libres créés spécifiquement pour un contexte médical vont être présentés.

LIBRAIRIES LOGICIELLES

Dans le domaine logiciel, il convient, avant toute chose, de distinguer les librairies et les applications. Les librairies logicielles sont des composants de code réutilisables conçus pour être intégrés dans d'autres logiciels, tandis que les applications sont des logiciels autonomes destinés à des tâches spécifiques.

Les librairies logicielles sont souvent considérées comme la «matière noire» de l'informatique parce qu'elles sont essentielles au fonctionnement des logiciels, mais sont généralement invisibles pour l'utilisateur final. En d'autres termes, elles fournissent des fonctionnalités de base et des outils pour le développement de logiciels, mais elles ne sont pas directement visibles dans l'interface utilisateur. À l'image de la matière noire en astronomie, les librairies sont omniprésentes, mais souvent peu comprises par les personnes qui ne développent pas des logiciels elles-mêmes.

STANDARDS D'INTEROPÉRABILITÉ

De nombreuses bibliothèques libres et open-source focalisées sur le domaine médical ont été développées. Cela s'explique, notamment, par l'existence de standards d'interopérabilité technique dédiés à la gestion de l'information médicale et adoptés à l'échelle internationale, tels que le HL7 (Health Level 7) et le FHIR («Fast Healthcare Interoperability Resources») (4). Le logiciel libre offre, en effet, l'avantage de mutualiser le développement d'implémentations de référence de ces standards, et de créer ainsi des infrastructures informatiques partagées, résilientes et peu onéreuses. En rassemblant une communauté de développeurs issus d'horizons différents autour d'un même standard, la création d'une implémentation mutualisée permet de réduire les coûts de développement et de diminuer la barrière à l'entrée, tout en augmentant la fiabilité, la couverture et la pérennité par rapport à une implémentation qui serait créée isolément par une firme unique. Une externalité capitale de l'existence d'une implémentation libre et open-source d'un standard d'interopérabilité médicale est qu'elle permet à la communauté scientifique de bénéficier des mêmes outils que ceux utilisés dans le milieu industriel, et ainsi de rapprocher la recherche médicale des systèmes d'information hospitaliers.

Concrètement, le logiciel Mirth Connect offre ainsi une implémentation libre et reconnue du standard HL7, qui vise à faciliter la communication entre les différents systèmes d'information de santé, notamment les dossiers médicaux électroniques, les systèmes de gestion administrative des patients, ou encore les systèmes de facturation. Mirth Connect est rédigé dans le langage Java. Les développeurs HL7 qui utilisent le langage Python recourent, quant à eux, fréquemment à la bibliothèque `python-hl7`. Le plus récent standard FHIR est, pour sa part, souvent décrit comme une évolution du HL7 qui modernise et simplifie l'échange d'informations de santé. Le FHIR repose sur les technologies Web (plus précisément, des API de type REST) et définit des «ressources» comme étant des éléments de données normalisés qui représentent différentes entités ou concepts dans le domaine des soins de santé. L'implémentation libre et open-source actuellement la plus connue du FHIR est dénommée HAPI et est rédigée en Java.

IMAGERIE MÉDICALE

Le DICOM («*Digital Imaging and Communications in Medicine*») est le standard international *de facto* pour l'imagerie médicale numérique (5). Utilisé à l'échelle mondiale depuis plusieurs

décennies, le DICOM autorise la gestion, le stockage et la transmission des images médicales et des informations associées. Il a été conçu pour garantir l'interopérabilité des systèmes d'imagerie médicale, permettant aux professionnels de la santé de partager facilement des images et des données, quels que soient le fabricant du matériel ou le logiciel utilisé. Le DICOM supporte non seulement toutes les modalités d'imagerie traditionnelles (radiographie, mammographie, échographie, CT-scanner, PET-scanner, IRM...), mais aussi les données utilisées en radiothérapie et protonthérapie (RT-STRUCT, RT-DOSE et RT-PLAN), les résultats d'algorithmes d'intelligence artificielle (DICOM-SR et DICOM-SEG), l'anatomopathologie, les signaux neurophysiologiques (EEG et ECG), ou encore les modèles 3D pour les implants ou la bio-impression (encapsulation de fichiers STL).

En fonction du langage de programmation utilisé, les développeurs informatiques pourront se tourner vers différentes implémentations libres et open-source du DICOM, telles que `dcm4che` pour le Java (6), `pydicom` pour le Python (7), ou bien `DCMTK` et `GDCM` pour le C++ (8). De nombreuses bibliothèques logicielles dédiées à l'affichage, à la manipulation et à l'analyse des images médicales ont également été développées, au premier rang desquelles VTK (9) et ITK (10), qui se retrouvent au cœur de nombreux logiciels. L'alignement d'images médicales est un domaine de recherche à part entière qui a mené au développement de nombreuses bibliothèques, particulièrement `Plastimatch` (11). Dans le domaine de la radiothérapie, `dicompyler` offre une plateforme de recherche dédiée à l'analyse des fichiers DICOM RT (12).

Terminons cette section en mentionnant le fait que la recherche en neuroimagerie recourt fréquemment au standard NIfTI («*Neuroimaging Informatics Technology Initiative*») (13). En effet, même si la neuroimagerie est parfaitement intégrée au DICOM, les équipes de recherche ne bénéficient pas toutes de la présence d'informaticiens capables de gérer la complexité du DICOM. Le NIfTI a, dès lors, été développé comme une alternative très simplifiée au DICOM, qui permet aux scientifiques de se focaliser uniquement sur le contenu des pixels ou voxels, sans devoir gérer les informations liées au patient ou à l'acquisition. Le standard émergent BIDS («*Brain Imaging Data Structure*») spécifie, quant à lui, comment organiser et documenter l'ensemble des données d'une étude en neuroimagerie (14). BIDS utilise le format NIfTI pour stocker les images, mais il va au-delà en normalisant la structure de répertoire et

les métadonnées. Des bibliothèques libres et open-source en Python, telles que NiBabel, offrent une prise en charge rapide du NIFTI.

INTELLIGENCE ARTIFICIELLE (IA)

Le logiciel libre et open-source a joué un rôle crucial dans l'avancement de l'IA, en particulier dans les domaines de l'apprentissage automatique («machine learning») et de l'apprentissage profond («deep learning»). La mise à disposition de bibliothèques puissantes pour l'IA a, en effet, permis de réduire les barrières à l'entrée pour de nombreux chercheurs et développeurs. Cette dynamique d'ouverture a également favorisé la collaboration et le partage de modèles, de données et d'algorithmes entre chercheurs du monde entier, accélérant ainsi grandement le progrès scientifique ces dernières années.

De nos jours, le langage de programmation le plus couramment utilisé pour l'intelligence artificielle, et plus largement pour la science des données, est le Python. Ceci s'explique par la grande variété de bibliothèques libres et open-source disponibles à cet usage dans l'environnement Python. Par exemple, les bibliothèques NumPy et SciPy permettent de manipuler efficacement des matrices et des algorithmes statistiques ou numériques (15). Le logiciel Pandas offre une solution très populaire pour rassembler différentes sources de données, et l'affichage de données est rendu possible par Matplotlib (16). Les chercheurs actifs en IA apprécient tout particulièrement la possibilité de créer des «notebooks» grâce à l'outil JupyterLab qui permet d'exécuter de manière interactive des analyses de données, puis de les partager avec d'autres chercheurs (17).

La réputation du Python découle également du fait que les trois bibliothèques les plus populaires dédiées à l'apprentissage profond sont immédiatement accessibles depuis ce langage. Il s'agit de TensorFlow (18), PyTorch (19) et Scikit-learn (20), qui sont toutes trois des bibliothèques libres et open-source. La bibliothèque Scikit-learn se distingue des deux autres car elle offre également des primitives pour l'apprentissage automatique classique. Le langage Python offre également un accès aisé à des bibliothèques spécialisées pour certaines catégories d'applications, telles que OpenCV pour la vision par ordinateur, ou SpaCy pour le traitement naturel du langage. En pratique, toutes ces possibilités sont, bien entendu, combinées avec les bibliothèques Python présentées plus haut pour créer de nombreuses applications innovantes dans le secteur médical.

APPLICATIONS

Le chapitre précédent a présenté une vue de haut niveau de quelques bibliothèques libres et open-source disponibles pour le secteur de la santé. Ces bibliothèques logicielles sont, avant tout, dédiées aux techniciens et restent, en général, invisibles pour l'utilisateur final. Mais à l'image des briques Lego, ces bibliothèques peuvent être combinées pour créer des applications de plus haut niveau accessibles à une audience médicale. Dans cette section, nous donnerons trois exemples de telles applications.

ORTHANC

Orthanc est une application libre et open-source qui est directement issue des recherches du Département de Physique médicale du CHU de Liège (21, 22). Il s'agit d'un serveur PACS («*Picture Archiving and Communication System*») qui permet de stocker, d'indexer et de communiquer les images médicales selon le standard DICOM. En interne, Orthanc est écrit en C++ et combine la bibliothèque DCMTK avec une base de données et avec un serveur Web embarqué.

Les trois principales spécificités d'Orthanc sont de publier une interface de programmation permettant d'automatiser des flux d'imagerie de manière sécurisée, y compris entre plusieurs sites, de proposer une architecture très légère qui permet le déploiement d'un grand nombre de serveurs Orthanc pour gérer de multiples flux au sein d'un même hôpital (on parle de micro-service), et d'offrir un système de plugins qui autorise l'extension du logiciel de base à de nouveaux usages. Grâce à ce système d'extensions, Orthanc a, notamment, intégré une des premières implémentations du standard DICOMweb, qui gouverne l'échange de données d'imagerie médicale à travers l'infrastructure d'Internet, ainsi que du DICOM pour l'anatomopathologie (23).

De nombreuses applications ont été développées sur base d'Orthanc, telles que des solutions de télédermatologie, des plateformes de routage automatisé avec anonymisation pour les études cliniques en lien avec l'industrie pharmaceutique, des outils d'enseignement de la radiologie, ou encore des passerelles vers des solutions d'IA hébergées dans le «cloud». Le projet Orthanc est un projet de grande ampleur qui porte une attention toute particulière à la qualité de son code. Orthanc a été récompensé, dès 2015, par la Free Software Foundation et son impact sociétal a récemment été reconnu

par le label «Digital Public Good» de l'Alliance pour les biens publics numériques (DPGA), une initiative multipartite approuvée par les Nations Unies qui «*facilite la découverte et le déploiement de technologies ouvertes, afin de créer un écosystème mondial prospère pour les communs numériques et ainsi contribuer à la réalisation des objectifs de développement durable*». (<https://digitalpublicgoods.net/>). Orthanc a ainsi été reconnu comme contribuant aux objectifs de développement durable numéros 3 (santé) et 9 (infrastructure et innovation).

Pour conclure, remarquons que des PACS libres et open-source autres que Orthanc existent, tels que Conquest, Dicoogle (24) et dcm4chee (6).

VISUALISATEURS POUR L'IMAGERIE MÉDICALE

Les visualisateurs d'images médicales au format DICOM forment une catégorie très importante d'applications. Il faut distinguer, d'une part, les visualisateurs «desktop» qui prennent la forme de logiciels de bureau et, d'autre part, les visualisateurs Web qui sont accessibles à travers un navigateur Internet. Dans les deux cas, des logiciels libres et open-source sont disponibles.

Un visualisateur desktop très populaire est 3D Slicer (25). Il est largement utilisé dans le domaine de la recherche médicale et offre des fonctionnalités avancées de visualisation, de segmentation et de traitement d'images volumiques. Tout comme Orthanc, 3D Slicer bénéficie d'un système d'extensions qui ajoutent

des fonctionnalités supplémentaires à l'application. Parmi ces extensions, on peut, par exemple, trouver SlicerRT qui est conçu pour la radiothérapie et qui peut prendre en charge la segmentation des structures anatomiques et la visualisation des plans de traitement (26). Une autre extension récente, dénommée TotalSegmentator, permet de segmenter automatiquement 104 structures anatomiques différentes grâce à l'apprentissage profond (27). Parmi les autres visualisateurs desktop libres et open-source disponibles, on peut également citer Horos et MedInria. La plupart de ces visualisateurs desktop s'appuient sur les bibliothèques VTK et ITK mentionnées plus haut.

En ce qui concerne les visualisateurs Web, ceux-ci sont actuellement en plein essor. Le plus connu est sans doute OHIF («*Open Health Imaging Foundation*»), qui offre des fonctionnalités comme les annotations, la segmentation du signal SUV pour la médecine nucléaire et les «*hanging protocols*», avec une focalisation sur l'oncologie (28). OHIF s'appuie sur la bibliothèque d'affichage d'images médicales pour le Web, dénommée Cornerstone3D, qui est rédigée en JavaScript. Kitware VolView est un autre visualisateur Web, qui est destiné à l'affichage interactif et cinématique de volumes 3D (29). Un troisième visualisateur Web s'appelle Stone Web Viewer, qui est directement issu des recherches autour d'Orthanc (30). Stone Web Viewer se distingue de OHIF et Kitware VolView par son approche adaptée à la téléradiologie en 2D plutôt qu'au rendu volumique (Figure 1). D'un point de vue technologique, Stone Web Viewer est presque

Figure 1. Stone Web Viewer affichant une IRM cérébrale (30)

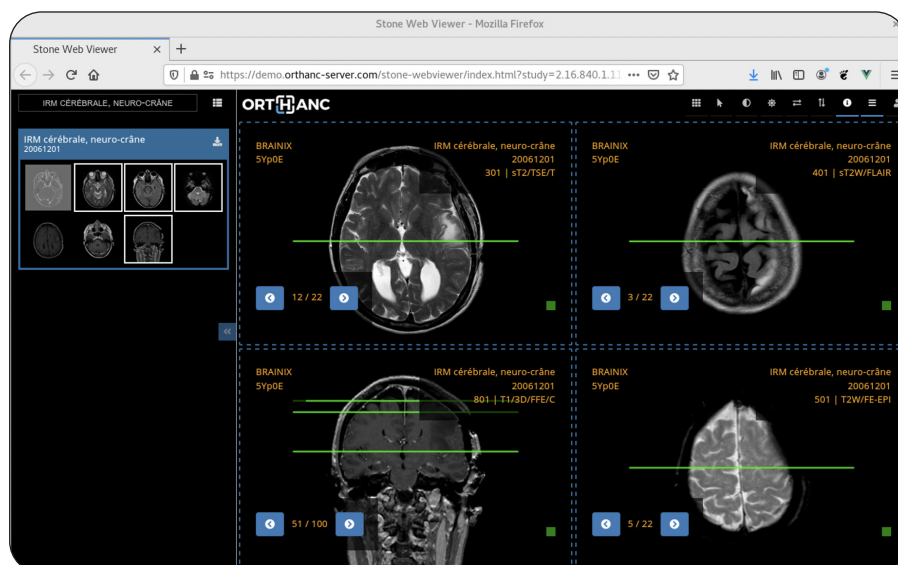
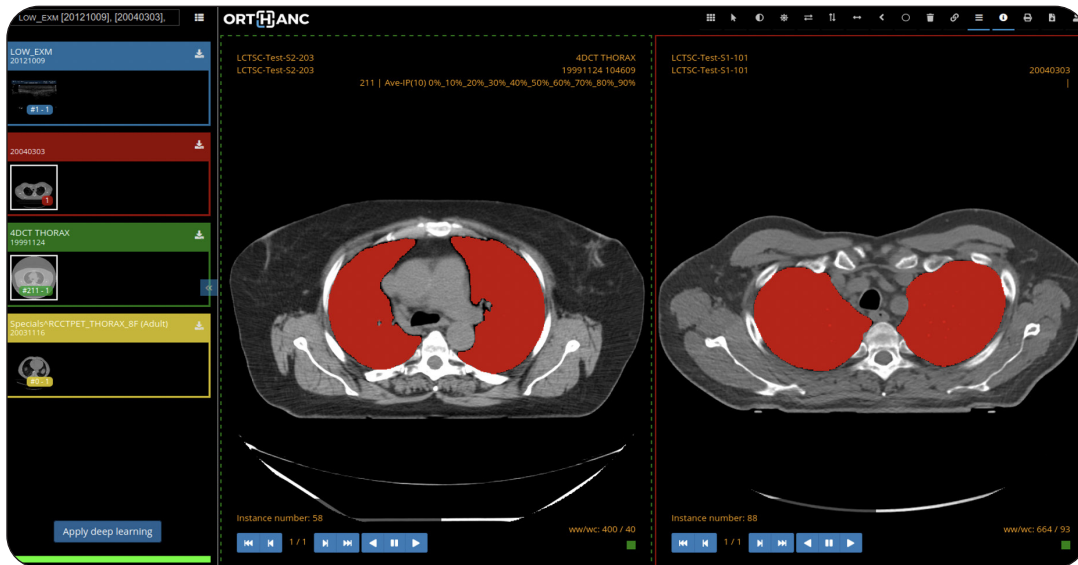


Figure 2. Segmentation automatisée du poumon par un modèle de deep learning à travers Stone Web Viewer (30)



entièrement développé en C++, grâce à la technologie émergente du WebAssembly, ce qui lui permet de partager une quantité importante de code source avec Orthanc. Il a aussi récemment été démontré que des algorithmes d'apprentissage profond peuvent être exécutés directement sur un navigateur Internet à travers Stone Web Viewer (Figure 2). La librairie libre Stone qui forme le cœur de Stone Web Viewer a, par ailleurs, été utilisée pour la création d'une application pour l'éducation thérapeutique du patient en radiothérapie (31).

Tous ces visualisateurs Web peuvent être utilisés à travers des navigateurs Internet libres et open-source comme Mozilla Firefox et Chromium. Remarquons, enfin, que les trois visualisateurs Web précités (OHIF, Kitware VolView et Stone Web Viewer) sont disponibles comme plugins pour Orthanc, ce qui facilite grandement leur installation par une audience non technique.

DOSSIERS PATIENTS INFORMATISÉS

Outre les applications en imagerie médicale et les implémentations de standards d'interopérabilité, des dossiers patients informatisés (DPI) libres et open-source sont également disponibles (32). Le projet GNU Health est un exemple d'un tel DPI. Ce projet vise à fournir un système de gestion hospitalière complet et intégré, ainsi qu'un système d'information pour les établissements et les professionnels de la santé. GNU Health est rédigé dans le langage Python, sur base de la plateforme applicative Tryton qui a été créée en Belgique. La vision de

GNU Health dépasse une dimension purement technologique, car ce projet porte une attention à la «médecine sociale», ce qui inclut, notamment, la promotion de la santé et la prévention médicale dans les régions rurales. GNU Health est développé par l'organisation non gouvernementale GNU Solidario et a, lui aussi, été reconnu comme Digital Public Good, à l'image d'Orthanc, en lien avec les objectifs de développement durable 3 (santé) et 11 (villes et communautés durables). Au niveau de l'imagerie médicale, GNU Health propose un module permettant de créer des liens entre le dossier médical d'un patient et les examens stockés dans un serveur Orthanc, avec un accès immédiat aux visualisateurs Web. D'autres DPI libres et open-source sont disponibles, tels que OpenEMR, GNUmed et FreeMedForms.

AUTRES EXEMPLES

Cet article ne se veut nullement exhaustif. Il faut souligner qu'un très grand nombre de bibliothèques et d'applications libres et open-source sont disponibles de nos jours dans tous les domaines de la médecine. Prenons, par exemple, Debian, qui est un système d'exploitation libre et open-source, célèbre pour sa stabilité et son engagement envers le logiciel libre, basé sur le noyau Linux et sur le GNU Project. Au sein de Debian, on retrouve le groupe Debian Med, qui se consacre à la distribution de nombreux logiciels liés au domaine de la santé, par exemple pour la biologie, l'épidémiologie, la pharmacologie ou encore l'oncologie.

L'équipe Debian Med gère plus de mille paquets (<https://blends.debian.org/med/tasks/>), dont une centaine rien que pour l'imagerie médicale. Même si la qualité des logiciels sous-jacents peut être variable et même si tous les logiciels existants sont loin d'être distribués dans Debian, ces chiffres illustrent l'impressionnante dynamique des logiciels libres dans le domaine médical.

Comme autres exemples d'applications libres et open-source, nous pouvons citer Cytomine qui est focalisé sur l'enseignement de l'anatomopathologie ainsi que sur l'analyse des images à très haute résolution par apprentissage automatique (33). Le projet OpenTPS est un «*Treatment Planning System*» (TPS) adapté à la recherche en radiothérapie et en protonthérapie par méthodes de Monte-Carlo (34). Bioconductor s'attache, quant à lui, à l'analyse statistique de données génomiques (35). Terminons en mentionnant MONAI, qui propose un environnement pour l'annotation d'images médicales et l'entraînement de modèles d'apprentissage automatique associés (36).

CONCLUSION

Du fait de la dépendance grandissante de la recherche scientifique vis-à-vis des algorithmes, il devient de plus en plus nécessaire de combiner les trois principes de l'open-access, de l'open-data et de l'open-source. Comme nous l'avons démontré, de nombreux logiciels libres et open-source sont aujourd'hui disponibles dans le secteur de la médecine, à la fois comme bibliothèques pour développeurs ou comme applications pour utilisateurs finaux. Au-delà de leur importance pour la recherche, les logiciels libres s'inscrivent dans une éthique de partage des connaissances techniques à une échelle internationale, ainsi que dans une optique de réduction des barrières technologiques, non seulement pour des applications industrielles et hospitalières, mais aussi vis-à-vis des économies émergentes dans une logique de coopération au développement.

Les logiciels libres vont très probablement jouer un rôle de plus en plus important pour l'implémentation de standards d'interopérabilité toujours plus évolués, pour la formation du personnel soignant en-dehors d'un hôpital, ainsi que pour l'empowerment du patient. La distribution à grande échelle de modèles d'intelligence artificielle robustes dédiés à une tâche médicale et accessibles à une audience généraliste sous une forme de logiciels libres est désormais un

enjeu essentiel dans le domaine de l'informatique médicale. En particulier, dans le contexte de l'imagerie médicale, le projet Orthanc souscrit pleinement à ces objectifs depuis sa création en 2011 grâce à l'impulsion du Département de Physique médicale du CHU de Liège.

BIBLIOGRAPHIE

1. Clark K, Vendt B, Smith K, et al. The Cancer Imaging Archive (TCIA): maintaining and operating a public information repository. *J Digit Imaging* 2013;**26**:1045-57.
2. Tomczak K, Czerwińska P, Wiznerowicz M. Review The Cancer Genome Atlas (TCGA): an immeasurable source of knowledge. *Współczesna Onkologia* 2015;**1A**:68-77.
3. Williams S. *Free as in freedom (2.0) : Richard Stallman and the free software revolution*. 2nd ed. Boston, Massachusetts:GNU Press;2010.
4. Benson T, Grieve G. *Principles of health interoperability, FHIR, HL7, and SNOMED CT*. 4th ed. New-York:Springer International Publishing;2021.
5. Kahn CE, Carrino JA, Flynn MJ, et al. DICOM and radiology: past, present, and future. *J Am Coll Radiol* 2007;**4**:652-7.
6. Warnock MJ, Toland C, Evans D, et al. Benefits of using the DCM4CHE DICOM archive. *J Digit Imaging* 2007;**20**:125-9.
7. Mason D. SU-e-t-33: Pydicom: an open source DICOM library. *Medical Physics* 2011;**38**:3493.
8. Eichelberg M, Riesmeier J, Wilkens T, et al. Ten years of medical imaging standardization and prototypical implementation: The DICOM standard and the OFFIS DICOM toolkit (DCMTK). Available from: <http://tinyurl.com/42y7ncyv>
9. Schroeder W, Martin K, Lorensen B. *The visualization toolkit, an object-oriented approach to 3D graphics*. 4th ed. New York: Kitware;2006.
10. McCormick M, Liu X, Jomier J, et al. ITK: Enabling reproducible research and open science. *Frontiers in Neuroinformatics* 2014;**8**:13.
11. Shackelford J, Kandasamy N, Sharp G. Plastimatch-an open-source software for radiotherapy imaging. In: *High performance deformable image registration algorithms for manycore processors*. 1st ed. Amsterdam;Elsevier:2013. p107-14.
12. Panchal A, Keyes R. SU-GG-t-260: dicompyler: an open source radiation therapy research platform with a plugin architecture. *Medical Physics* 2010;**37**:3245.
13. Li X, Morgan PS, Ashburner J, et al. The first step for neuroimaging data analysis: DICOM to NIfTI conversion. *J Neurosci Methods* 2016;**264**:47-56.
14. Gorgolewski K, Auer T, Calhoun V, et al. The brain imaging data structure, a format for organizing and describing outputs of neuroimaging experiments. *Sci Data* 2016;**3**:160044.
15. Virtanen P, Gommers R, Oliphant TE, et al. SciPy 1.0: fundamental algorithms for scientific computing in Python. *Nat Methods* 2020;**17**:261-72.
16. McKinney W. *Python for data analysis: data wrangling with pandas, numpy, and ipython*. 2nd ed. Sebastopol;O'Reilly Media:2017.
17. Randles BM, Pasquetto IV, Golshan MS, Borgman CL. Using the Jupyter notebook as a tool for open science: an empirical study. *ACM/IEEE Joint Conference on Digital Libraries (JCDL)* 2017:1-2. Available from: <https://ieeexplore.ieee.org/document/7991618>
18. Abadi M, Barham P, Chen J, et al. TensorFlow: a system for large-scale machine learning. In 12th Usenix Symposium on Operating Systems Design and Implementation (OSDI 16), Nov. 2016, pp. 265-283. Available from: <https://www.usenix.org/conference/osdi16/technical-sessions/presentation/abadi>

19. Ketkar N. Introduction to PyTorch. *In deep learning with python*. Berkeley;Apress:2017. p195-208.
20. Pedregosa F, Varoquaux G, Gramfort A, et al., Scikit-learn: machine learning in Python. *J Mach Learn Res* 2011;**12**: 2825-30.
21. Jodogne S, Bernard C, Devillers M, E, et al. Orthanc: a lightweight, RESTful DICOM server for healthcare and medical research. Proceedings, IEEE International Symposium on Biomedical Imaging: from Nano to Macro, pp. 190-193, 2013. Available from: <https://ieeexplore.ieee.org/document/6556444>
22. Jodogne S. The Orthanc ecosystem for medical imaging. *J Digitl Imaging* 2018;**31**:341-52.
23. Jodogne S, Lenaerts E, Marquet L, et al. Open implementation of DICOM for whole-slide microscopic imaging. In VISIGRAPP 2017 conference, pp. 81–7. Available from: <https://www.scitepress.org/Link.aspx?doi=10.5220/0006155100810087>
24. Costa C, Ferreira C, Bastião L, et al. Dicoogle: an open-source peer-to-peer PACS. *J Digitl Imaging* 2010;**24**:848-56.
25. Pieper S, Halle M, Kikinis R. 3D Slicer. In 2nd IEEE International Symposium on Biomedical Imaging: Macro to nano (IEEE cat no. 04EX821), Conference 2004. Available from: <https://ieeexplore.ieee.org/document/1398617>
26. Pinter C, Lasso A, Wang A, et al. Slicer RT, radiation therapy research toolkit for 3D Slicer. *Med Phys* 2012;**39**:6332-8.
27. Wasserthal J, Breit HC, Meyer MT, et al. Total Segmentator: Robust segmentation of 104 anatomic structures in CT images. *Radiol Artif Intell* 2023;**5**:e230024.
28. Ziegler E, Urban T, Brown D, et al. Open health imaging foundation viewer: an extensible open-source framework for building web-based imaging applications to support cancer research. *JCO Clinl Cancer Inform* 2020;**4**:336-45.
29. Xu J, Thevenon G, Chabat T, et al. Interactive, in-browser cinematic volume rendering of medical images. *Comput Methods Biomech Biomed Eng Imaging Vis* 2022;**11**:1019-26.
30. Jodogne S. On the use of WebAssembly for rendering and segmenting medical images. In Roque AC, Gracamin D, Lorenz R, editors: *Biomedical engineering systems and technologies*. Cham;Springer Nature Switzerland:2023. p393-414.
31. Kirkove D, Barthelemy N, Coucke Ph, et al. Étude de faisabilité : l'utilisation de l'imagerie médicale en éducation thérapeutique en radiothérapie. *Cancer Radiother* 2022;**26**:1034-44.
32. Syzdykova A, Malta A, Zolfo M, et al. Open-source electronic health record systems for low-resource settings: systematic review. *JMIR Med Inform* 2017;**5**:e44.
33. Marée R, Rollus L, Stévens B, et al. Collaborative analysis of multi-gigapixel imaging data using cytomine. *Bioinformatics* 2016;**32**:1395-1401.
34. Wuycens S, Dasnoy D, Janssens G, et al. OpenTPS: Open-source treatment planning system for research in proton therapy. arXiv, 2023. Available from: <https://arxiv.org/abs/2303.00365>
35. Gentleman RC, Carey VJ, Bates DM, et al. BioConductor: open software development for computational biology and bioinformatics. *Genome Biol* 2004;**5**:R80.
36. Cardoso MJ, Li W, Brown R, et al. MONAI: an open-source framework for deep learning in healthcare. arXiv, 2022. Available from: <https://arxiv.org/abs/2211.02701>

Les demandes de tirés à part doivent être adressées au Pr Jodogne S, ICTEAM, UCLouvain, Louvain-La-Neuve, Belgique.

Email : sebastien.jodogne@uclouvain.be