✕USCAP

ARTICLE

Digital media archive for gross pathology images based on open-source tools and Fast Healthcare Interoperability Resources (FHIR)

Emilio Madrigal [iD][1] · Long Phi Le [iD][1]

Abstract

Macroscopic examination of surgical pathology and autopsy cases is a fundamental component of anatomic pathology. The photographic documentation of such clinical specimens is essential, and it may be required in certain instances. Our department began using consumer-grade digital cameras in 2005 to improve the practice of gross photography. However, the lack of an application to correctly catalog the photographs resulted in thousands of digital image files scattered across shared network drives, with limited case and patient metadata, making image retrieval a cumbersome and sometimes impossible task. Thirteen years later, we examined the legacy method of acquiring and accessing gross photographs in our department and determined the need for a web-based digital media archive to capture images with structured metadata. Using several open-source tools, including MediaWiki, we developed a flexible platform for building our digital media archive with a data schema based on the Fast Healthcare Interoperability Resources standard. Following a short pilot, we replaced the legacy method of handling gross pathology images with a new acquisition workflow and digital media archive. Through March 2021, 233 distinct users have accessed the system, 58 of which have uploaded 21,024 images. Of those images, 13,684 (65.1%) correspond to surgical pathology images, 4045 (19.2%) belong to neuropathology cases, and 3295 (15.7%) originate from autopsies. We demonstrate the design and implementation of a customizable anatomic pathology digital media archive solution in an academic pathology department setting using a modern standard for exchanging healthcare information electronically. The system's efficiency and scalability for our current operation will enable us to integrate with other applications and pathology informatics initiatives in the future.

Introduction

The digitization of photographic prints and 35 mm slides enabled by the introduction of digital scanning systems in the mid-1980s, along with the availability of consumer-grade digital cameras in the early 1990s, set off the early adoption of digital pathology in anatomic pathology (AP) laboratories [1]. Since then, pathology departments have generated countless digital files, including gross photographs, to document macroscopic findings for surgical specimens and autopsies. These photographs are essential to the practice of AP and pathology education, and are

diagnostic observations which belong in the electronic health record [2, 3].

Over the last three decades, digital cameras have become ubiquitous in autopsy suites and gross stations; however, logistical and technical challenges have deterred the wide-spread adoption of digital media archives (DMAs). Without an application to manage the creation, indexing, workflow, version, and access control of files, AP laboratories have resorted to implementing imperfect workarounds to store and retrieve digital images [4–7]. The Massachusetts General Hospital Department of Pathology began to routinely use digital photography in 2005 and has relied on shared network drives (SNDs) to manage its digital image files without standardized naming structures or accessible metadata.

This article describes the development and implementation of a web-based DMA for gross photographs and other image files at a major academic pathology department. We built the system using open-source tools to overcome multiple deficiencies in our legacy solution, including

✉ Emilio Madrigal
  emadrigal@mgh.harvard.edu

[1] Department of Pathology, Massachusetts General Hospital,
  Harvard Medical School, Boston, MA, USA

limited accessibility and lack of structured data. The DMA schema is based on the Fast Healthcare Interoperability Resources (FHIR) standard.

Materials and methods

This quality improvement project was conducted at the Massachusetts General Hospital Department of Pathology, which in 2019 examined 78,000 surgical pathology cases and performed 435 autopsies. Since 2005, our department has used SNDs to archive and retrieve digital photographs; archived image files were analyzed with a Python 3.6.7 script. In the fall of 2018, we assessed the state of our legacy gross photography solution. We determined that a web-based DMA would notably improve the access, accuracy, efficiency, and value of managing our digital image files. We consulted pathologists' assistants (PAs) and technical staff to analyze the workflow of the department's traditional handling of gross image files (Fig. 1A). The hardware and network infrastructure were evaluated at all gross photography stations, including the main gross room, autopsy and neuropathology suites, and an affiliated neuropathology research facility.

Understanding the acquisition workflow and required features for our DMA, we developed a web application based on the MediaWiki 1.31.1 content management system driven by PHP 7.2.11 and MariaDB 10.4.2 (Table 1) [8]. The system was deployed with Docker images on a Rancher v2.4.8 cluster that hosts Kubernetes v1.18.6 and consists of three master nodes and four worker nodes all running locally (Fig. 2). The official MediaWiki 1.31.1 Docker image is based on the Debian 9 Linux operating system and was modified with additional libraries and utilities, including the PHP library to support LDAP authentication. All Docker images were pulled from their official repositories in Docker Hub. The worker nodes that host the Docker containers are based on the Ubuntu 18.04.5 LTS operating system. Application and configuration files are persisted locally on the worker node and mounted inside the running Docker container. The uploaded images are stored on a network file storage and the MariaDB database files are persisted locally on its worker node. The image and database files are fully backed up every hour during the business day to a remote network storage server. Three days' worth of backup database dumps are retained on a rolling basis.

Several open-source extensions drive the front-end and back-end functionality. The Cargo extension was used to declare the data schema for the content stored in structured query language (SQL) tables for query display on pages. The LDAP Authentication extension controls user authentication based on our enterprise Microsoft Active Directory with restricted access behind our institutional firewall.

Form-based data entry and presentation were achieved with the Page Forms extension. The SimpleBatchUpload extension provides a key interface for drag and drop image uploads, while the MultimediaViewer extension offers image viewing. Many forms, static pages, and templates support importing and retrieving gross images in the DMA. An auxiliary table was created for Systematized Nomenclature of Medicine Clinical Terms (SNOMED CT) body structure, physical object, and procedure codes to provide the values for the *reasonCode* and *bodySite* fields in the FHIR Media Resource. Additionally, we generate a daily extract from our AP laboratory information system (AP LIS; Sunquest CoPathPlus v6.3.2008) to match part designators with corresponding specimen descriptions. The extract was initiated with specimen data starting in 2017. The daily extract is imported into a custom SQL table to later join with the Media table, enriching the file metadata and improving image indexes. Custom functions were built to enable users to search for images using various parameters.

A prototype of the platform was introduced to a head PA for testing. Selected archived images were added to the system retrospectively. After addressing reported issues and requested enhancements, an updated version was released to additional PAs and residents. We used a questionnaire as a user acceptance tool, which resulted in other suggestions incorporated into a final stable release. The system was introduced to all departmental members in email communications, which included a system video tutorial and user guide. We also demonstrated the system in person at faculty and resident meetings. For a little over a year, the application ran on a custom SQL data schema to capture specimen metadata and user upload details. More recently, we switched to the FHIR v4.0.1 standard to leverage a common data schema. We converted and migrated our custom data structure to FHIR by relying on the annotation and status data types and the Media Resource [9]. Data migration involved a Python 3.6.7 script using a MediaWiki application programming interface (API) client (mwclient-0.10.1) with logging to capture time metrics.

Results

After our department adopted digital photography in 2005 using the retired legacy workflow, two technical staff members had archived 144,053 digital image files into an SND, accounting for 372 gigabytes of disk space. The drive contains 69,941 surgical pathology gross images with an annual average of 4662 files for specimens accessioned between 2005 and 2019. Autopsy images comprised 59,801 files with an annual average of 4600 files for cases accessioned between 2007 and 2019. A total of 5766 (8.2%)
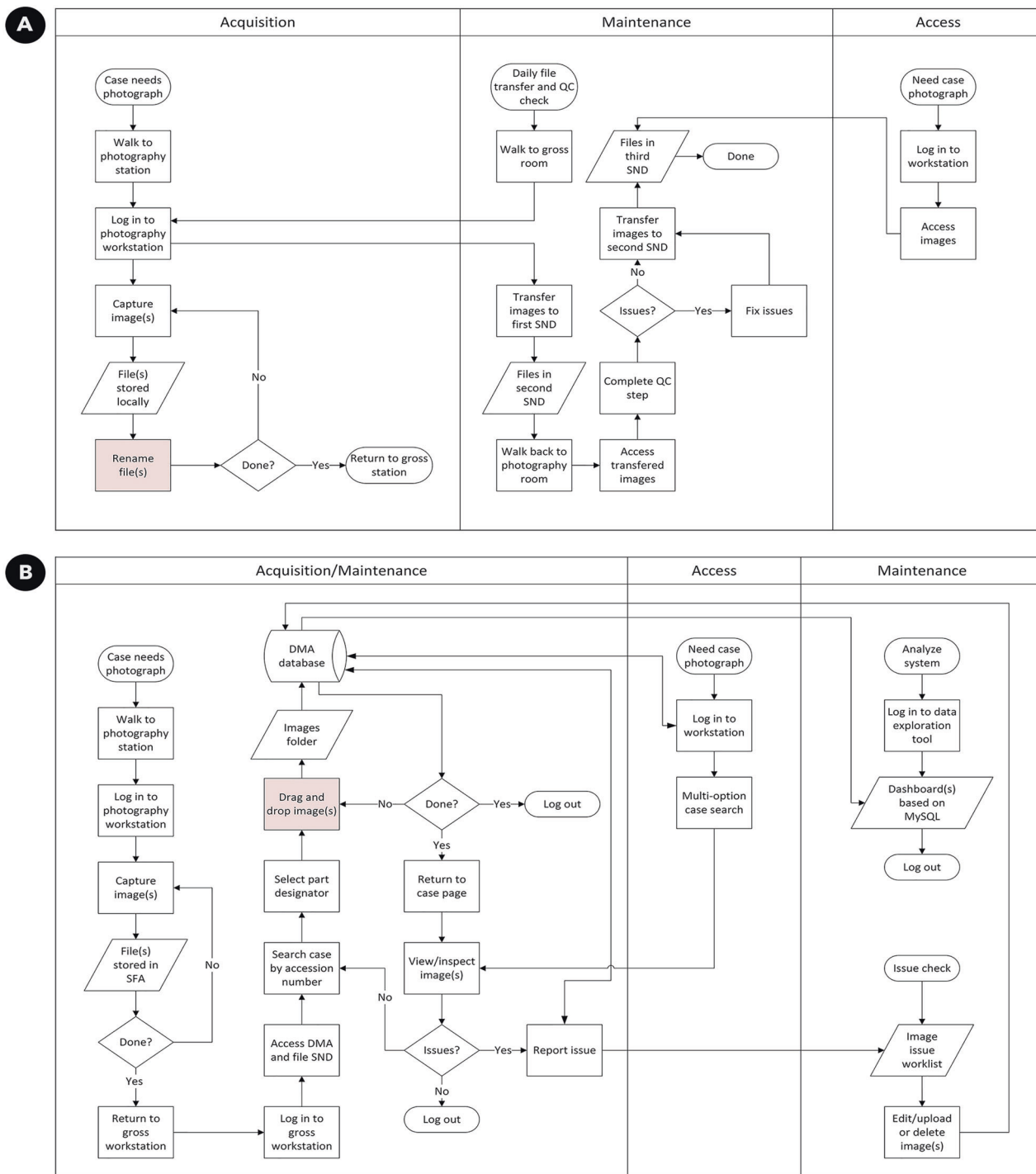
Fig. 1 Cross-functional flowcharts of legacy and current methods of handling surgical pathology gross photographs. A The legacy method involved a manual and error-prone step (red process box) resulting in labeling errors and inconsistent naming conventions with limited metadata, precluding the ability to search for images. B In the current method, the cataloging function is part of the ingestion process indicated by the red process box, reducing labeling errors and automatically associating rich metadata to the image file.

archived images were uploaded to the DMA, 399 for 2019, 3880 for 2018, and 1487 for 2017.

The DMA was developed over the course of 6 months with an estimated equivalent of one full-time engineer. The system supports three independent modules for surgical pathology, neuropathology, and autopsy images. Each module comprises one static page, three forms, two templates, and dozens of auxiliary templates. All modules share templates that define the database table schema for FHIR data types and the Media Resource. The Media Resource

Table 1 System components and essential extensions.

System software

MediaWiki 1.31.1

PHP 7.2.11

MariaDB 10.4.2

MediaWiki extensions

Name	Purpose
Cargo	Structured data entry and retrieval via SQL
Page forms	Javascript forms-driven data entry and query
SimpleBatchUpload	Drag-and-drop form for multiple file upload
LDAP authentication plugin	User creation and authentication based on LDAP
MultimediaViewer	Built-in media viewer for uploaded images
PrivatePageProtection	Read/write content restriction
UserFunctions	User/group-based content restriction

LDAP Lightweight directory access protocol, *PHP* PHP: hypertext preprocessor, *SQL* structured query language.
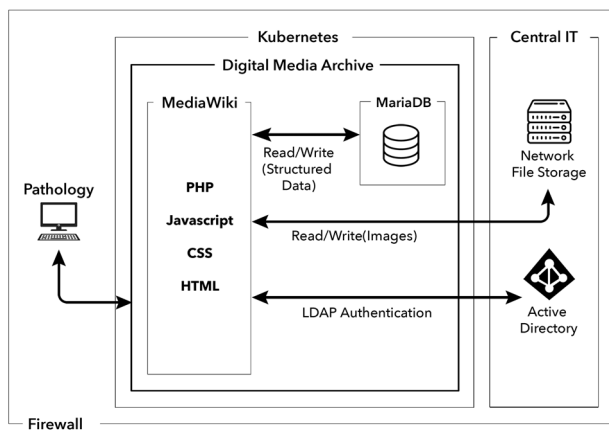

Fig. 2 System architecture. The digital media archive is powered by MediaWiki and MariaDB Docker images orchestrated by Kubernetes. MediaWiki provides the PHP code to generate the HTML, javascript, and CSS on content pages. MediaWiki interfaces with structured data in the MariaDB database and image files in a network file system. The entire application stack resides behind our institutional firewall and communicates with our Active Directory for LDAP authentication.

provides a mechanism for storing image files associated with a case accession identifier and other metadata. The actual media content is accessed by direct retrieval in an image viewer that is part of the built-in MediaWiki media content management. The annotation table handles the case part designator associated with an image. Issues (e.g., the need to delete, edit, reassign, and other functions) related to an image are flagged and designated in the FHIR annotation and status tables. A daily worklist is generated for a system administrator based on the image flags. When it was first

launched in January 2020, the system was designed to use a custom data schema to support case page creation and image uploads. After adopting the FHIR standard, the system was migrated to a new data schema. The old specimen and image metadata were converted to the new FHIR syntax by updating the respective MediaWiki file pages. Only the metadata had to be converted while the built-in MediaWiki handling of the file content was left intact. The migration execution time for the different image types was logged: 36 min for 12,591 surgical gross images, 30 min for 3405 neuropathology images, and 8 min for 2706 autopsy images. In addition, a MediaWiki maintenance script was executed in parallel for ~1 h to refresh the page contents and links after the FHIR standard update.

We designed the web-based DMA and the new image acquisition workflow to improve efficiency and to minimize file labeling errors (Fig. 1B). For surgical pathology and neuropathology images, the person handling the case transports the specimen to a central photography station near all gross stations. At the photography station, images are captured with a digital single-lens reflex camera, automatically named with a unique filename, and saved to an SND by the camera software. Autopsy images are captured similarly and frequently produced with a handheld camera, which requires manual transfer of files from the memory card to the SND.

After capturing the images, users access the DMA with the Google Chrome browser and navigate to the SND containing their images from a departmental workstation. Each user is authenticated in the DMA against our Microsoft Active Directory using a Lightweight Directory Access Protocol (LDAP). Following system login, users are presented with different home pages based on their assigned user rights (Fig. 3A). Access control to the system is based on six groups affecting a user's ability to perform specific actions. The process of uploading an image starts with completing a regular expression-controlled form field with the specimen or autopsy accession number, either typed or scanned using a barcode reader. The field ensures that the accession number format matches possible cases in our AP LIS. Submitting the form executes a query in the archive to find image files associated with the searched accession number in the FHIR Media table and joins the images with their respective part designators from the annotation table. An additional query searches for the case in our AP LIS extract table and determines all possible part designators for the case. If the specimen is associated with at least one photograph, a table in the web interface is presented with the image thumbnail, operator, upload date, and an action button to report issues with the image. The user interface shows an upload button for any queried specimen, which cascades into one or multiple buttons for all possible case part designators (Fig. 3B). If the case is not found in our AP
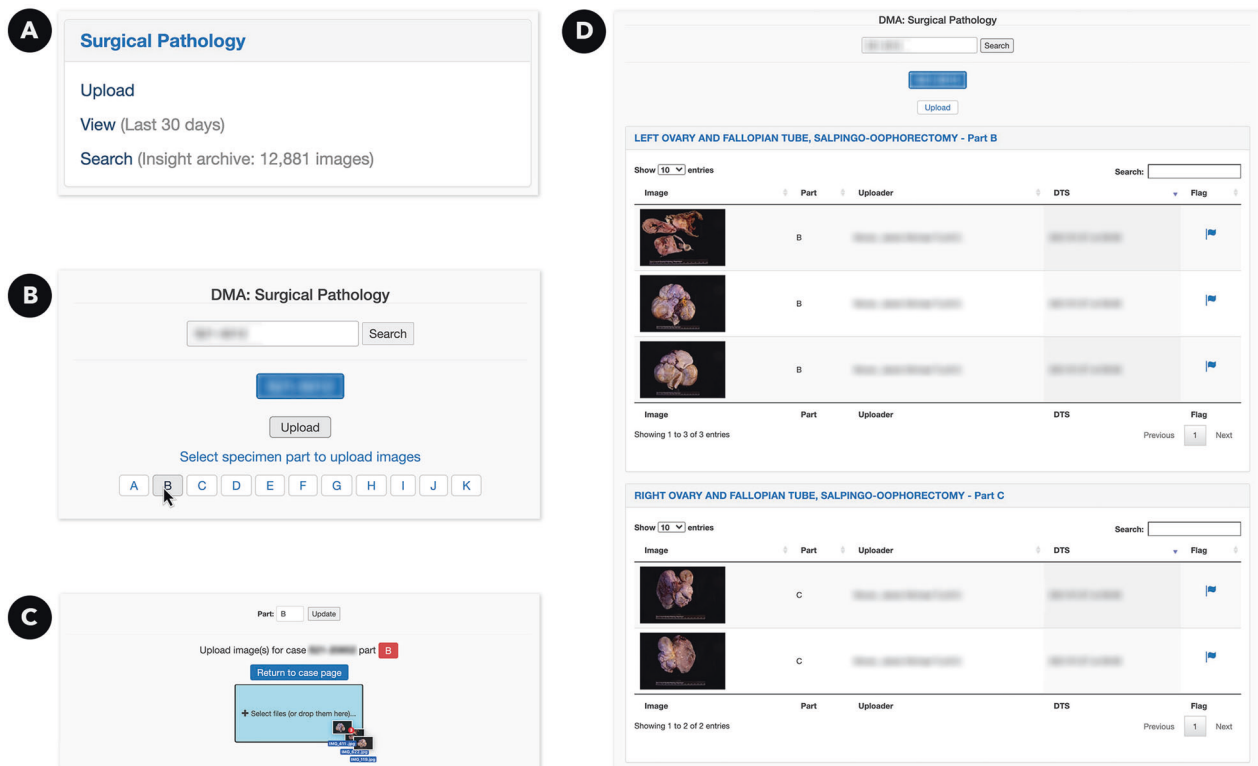
Fig. 3 Screen captures of the surgical pathology DMA's graphical user interface. A Users with DMA access are presented with home page elements enabling them to upload new images or search for uploaded files. **B** Once a user searches for a case accession number using a regular expression-controlled field, the system queries our laboratory information system extract and presents associated specimen designators. **C** Selecting the appropriate part enables a user to upload one or more images with a drag-and-drop action. **D** Case images are grouped by part and presented in vertically collapsing accordions.

LIS extract, the user sees eight action buttons representing eight generic part designators A-G along with a custom option, enabling users to define additional designators. For autopsies, users choose one of the 11 most commonly photographed body sites matched to our custom SNOMED CT codes table instead of part designators. If the image contains a different body site, the user can select the "Other" option, which presents them with a drop-down menu containing additional terms. Neuropathology cases do not require choosing a body site since a default term is assigned. As part of this research module, the user is presented with the option to select the autopsy type and disease category.

Selecting a part designator or body site executes a second query. If images already exist in the archive for the part or body site, they are presented to the user. A form represented as a drag and drop area is made available to the user for image upload. This is a vital feature of the DMA since it enables the cataloging function to occur automatically and simultaneously during the ingestion process of a new asset (Fig. 3C). By "dropping" one or more image files into the form area, users upload photographs to our image repository and automatically associate images with a case and its part designator or body site. The association is achieved by inserting a row per file in the FHIR Media table with nine fields, including an identifier, the subject (case accession number), operator identifier, and a date-time stamp, among others. During upload, the files are prefixed with the case accession number to avoid filename collision and guarantee a standard naming convention by avoiding manual renaming. The prefixed filename becomes the FHIR identifier. Uploading takes a couple of seconds per image, after which users are presented with a line of green-highlighted text indicating success along with the original and prefixed filenames. Any upload failure would show red-highlighted text instead. Once uploaded, the user can upload other images, change the part designator or body site, return to the case page to see all associated images (Fig. 3D), or navigate to upload images for a new case. If a user inadvertently uploads the wrong photograph, they may flag the image using an action button to indicate if the file needs to be deleted or reassigned. General users do not have access to modify uploaded files; therefore, delete and reassign requests are handled by a system administrator. Every action, including edit, delete, move, and upload, is logged in the system, creating an audit trail for every file.

Users have different options to access uploaded images through a web interface. Images for surgical pathology

cases are presented in a tabular format arranged by descending order of the upload date and time. The table shows the case accession number as a hyperlink leading to a page with all uploaded images, a count of uploaded files for the case, and the latest upload date. Users can filter the table using a built-in search function and sort the data by any of the columns. Additionally, users can use a custom search function to find cases based on accession number, upload user, or specimen designator. Since our AP LIS stores diagnoses for multiple parts and specimens as a single text blob and not as structured data, we cannot at this time associate discrete diagnosis information with images to drive queries. Neuropathology and autopsy images are presented using a similar dynamic table. Neuropathology images are enriched with autopsy type and disease category; hence users can search on three different autopsy types (Alzheimer's disease, adult, perinatal) and eleven disease categories. Rather than sorting through thousands of image files in the legacy solution (some of which were improperly named), users can now efficiently search for gross images using the DMA. For clinical signout, the updated workflow includes searching for all autopsy and neuropathology cases in the DMA. For surgical pathology cases, if the gross description indicates that an image was captured, or if a pathologist would like to review a gross photograph, they access images using one of the methods above.

Before production launch, archived images were uploaded to our web-based DMA in the summer of 2019. Starting in January 2020, the system was deployed to support prospective uploading of all newly generated photographs; the legacy system and workflow relying on simple storage on SNDs were retired. Through March 2021, 233 distinct users have accessed the system, 58 of which have uploaded a total of 21,024 images. Of these images, 13,684 (65.1%) correspond to surgical pathology images, 4045 (19.2%) belong to neuropathology cases, and 3295 (15.7%) originate from autopsies (Table 2). Shortly after launch, the World Health Organization declared COVID-19 a global pandemic, which

resulted in lower specimen volumes in our department, hence a decrease in system users and acquired photographs (Fig. 4). Nevertheless, the number of uploaded surgical pathology images for 2020 was comparable to previous years, which we attribute to the system's ease of use. The service with the most uploaded images was gastrointestinal pathology, with 2231 images, while the most uploaded specimen type was placenta (Fig. 5). In aggregate, the image files take up 48.57 gigabytes of disk space, with each ranging from 48 kilobytes to 10.2 megabytes. The images have averages of 2915 pixel width and 2059 pixel height. Most images were of jpeg type (21,003, 99.9%); 21 png and two pdf files make up the remainder. Users have reported a total of 275 issues, including 155 requests for images to be removed and 120 requests to reassign images to a different part, body site, or case.

Discussion

Digital imaging has significantly changed the practice of pathology around documentation of gross pathology findings [10, 11]. It began with the technology to convert photographic prints and 35 mm slides to digital images using scanners [1]. The availability of consumer-grade digital cameras quickly disrupted this practice, obviating the need to work with film and development in the darkroom while empowering users with instant access to digital results that could easily be transferred and shared. Like many other institutions, our department adopted both of the above transitions to digital media in pathology starting in 2005 and had amassed hundreds of thousands of images. While the ease of image generation, readily available equipment, and ever-decreasing storage costs lowered the barrier for us to embrace digital media in our practice, the technology to manage the deluge of daily content lagged far behind. Even today, vendor AP LISs fail to offer a complete solution to manage gross pathology images, requiring departments to build custom tools to meet their needs [5, 12]. Unlike other solutions, our system links images at the level of a case's specimen rather than at the whole case level. This specimen-level association method can serve as a model for other image types, including electron microscopy, fluorescence, and whole slide images. It also enables queries to be grouped by distinct specimen types without presenting all other unrelated specimen images of a case.

Our current AP LIS does not have a built-in DMA; instead, a file attachment activity enables users to copy and paste individual image file paths to a shared network drive. There are several disadvantages of this approach, including the laborious and error-prone process of manually linking file paths and the limited storage capacity of institutional SNDs. Our department previously used the file attachment

Table 2 Digital media archive analysis.

| Component	Observation
Images, total	21,024
Images, surgical pathology	13,684 (65.1%)
Images, neuropathology	4045 (19.2%)
Images, autopsy	3295 (15.7%)
Size, total	48.57 Gigabytes
Size, surgical pathology	26.2 Gigabytes
Size, neuropathology	12.23 Gigabytes
Size, autopsy	10.13 Gigabytes
System users, count	233
Users who uploaded images, count	58
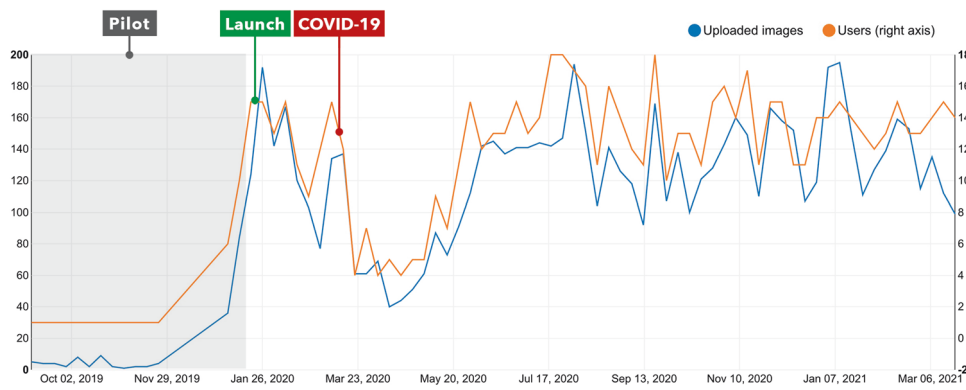
Fig. 4 Weekly volume of surgical pathology images and DMA users. The system was officially launched on January 23, 2020. Less than 2 months later, on March 11, 2020, the WHO declared COVID-19 a global pandemic, which eventually led to a decline in surgical procedures. The decrease in surgical pathology specimens is visually apparent in the steep decline of uploaded images (blue line, left Y-axis) and system users (orange line, right Y-axis).

Fig. 5 Volume of surgical pathology gross photographs by service and specimen type.
A After combining sub-services into main subspecialty groups, most of the uploaded surgical pathology images (2109, 16.4%) were accessioned to one of our department's four gastrointestinal pathology services. **B** The most uploaded specimen type is placenta, as illustrated by this word cloud, highlighting the variability in gastrointestinal pathology specimens compared to obstetric pathology.
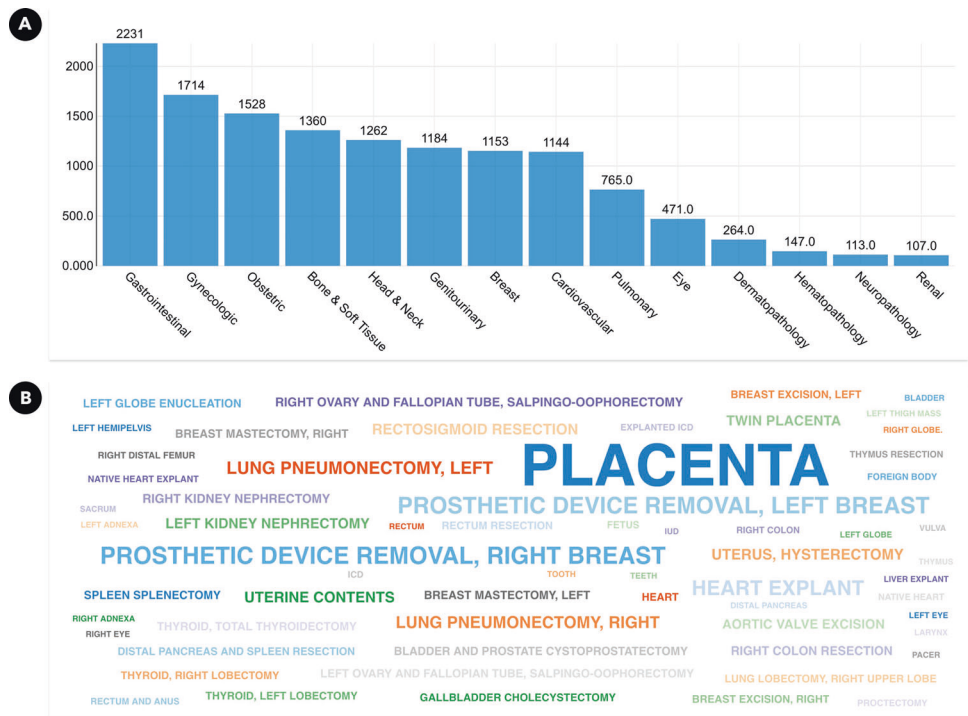
activity and realized another shortcoming when our institution changed directory paths which broke all previously entered image links. While we understand that vendors may improve the attachment activity's functionality, there are additional drawbacks with most enterprise pathology applications. Our AP LIS does not have an API for read or write functionality; there is no method to programmatically import data into or export data out of the system without significant work from the vendor to create an interface. We have prior experience working with the same vendor to create a whole slide imaging interface which evolved into a costly and time-consuming effort. Considering that the AP

LIS will change soon at our institution, the urgency to address the deficiencies of our legacy solution effectively ruled out a vendor solution. We explored the support of gross images in our new candidate AP LIS. Unfortunately, that application links images at the case level without more granular relation at the part level and lacks the function to include additional metadata.

Like many other institutions, our department's legacy image management system involved storing image files in an SND. The old system required one full-time employee to manage the files, complete quality control steps, and ensure file naming consistency. This was a laborious process

requiring image transfers, processing, and copying to a final destination folder. In addition to a notable delay in image availability, the system relied on folder and file navigation for retrieval. Folder management and file naming can only go so far in facilitating the organization and retrieval of images. As widely reported in the literature, manual data entry is an error-prone process that can be significantly improved with digital automation [13–15]. Even though a department-wide standardized file naming convention was in place, the daily volume of images made it impractical to police filenaming consistency, ultimately resulting in the haphazard reliance on each user's personal preference. The filename strings were at times formatted to include multiple unstructured data elements such as an accession number, part designator, initials of the operator, and information about specimen orientation. While informative at the single image level, this unstructured file naming practice fails at the system level where individual data elements are not accessible to search across many images. Indeed, our legacy SND archive of hundreds of thousands of images is worthless from the perspective of image search since no structured metadata was systematically captured.

Prompted by the lack of a customizable DMA solution from our AP LIS and other vendors, we set forth to develop our application with the following requirements in mind: (1) web-based application using open-source tools, (2) capture of structured metadata, (3) distribution of the image upload task at the point of specimen handling, (4) ease of use for the uploading step, and (5) flexibility to add or change both the back-end and front-end of the software. To meet these requirements, we selected the MediaWiki platform, an open-source content management system developed, supported, and maintained since 2001 [16]. One year after its initial development, MediaWiki became the engine for Wikipedia, the largest online crowd-sourced encyclopedia and ranked as the world's 13th most popular Alexa website [17]. MediaWiki as a content management system offers several desirable out-of-the-box features, including user management, content versioning/history, concurrent transaction handling for collaborative content creation/editing, and the ability to customize the user interface. Extending the basic functionalities of MediaWiki is possible through the installation of available extensions. We leveraged several extensions for our DMA to add LDAP authentication, access control, structured data management, forms-driven content editing, and easy drag and drop file uploads (Table 1).

Our initial laboratory-developed data schema was serviceable for over a year, but as patients move throughout our healthcare ecosystem, we anticipate the need to provide additional functionality and interoperate with other systems. For this purpose, we chose FHIR, which is a standard for exchanging healthcare information electronically, adopted by Health Level Seven International (HL7) [18]. Although FHIR has been described as a mechanism to support whole slide imaging, using the standard for gross images remains unexplored [19]. Applying the FHIR Media resource simplified our DMA implementation by eliminating three single-use, system-specific SQL tables, and creating three new multipurpose tables with data types that define common reusable patterns of elements. Our DMA is not limited to gross photography, and in fact, it can be extended to support various media types. In a future version of the DMA, we plan to add support for microscopy photomicrographs, electron microscopy images, immunofluorescence images, and histology block radiological images. Our current implementation uses FHIR as a stand-alone data exchange standard, but the potential remains for it to be used in concert with other standards; we will explore Digital Imaging and Communications in Medicine (DICOM) conversion for our images [12, 19].

The MediaWiki platform has allowed us to meet the above-listed requirements and address the need to have a flexible system. In addition to the gross pathology DMA, the platform's flexibility has enabled us to deploy other clinical applications in the same system, including histology slide requests, intraoperative/final diagnosis correlation, clinical pathology on-call logs, and SARS-CoV-2 testing. With the numerous functionalities that we have built into the platform, it has become our adjunct LIS to provide the much-needed customizability to supplement the deficiencies of our AP LIS. The flexibility of our platform was also demonstrated within the gross pathology DMA application itself. We initially developed the DMA to support gross clinical images for surgical pathology. Extending its capability to accommodate gross images for the autopsy clinical service and the neuropathology research program involved minor changes to the surgical gross pathology forms and templates to support custom metadata and display. We further demonstrated the platform's flexibility by overhauling our custom data schema. Several forms and templates were developed based on the FHIR annotation and status data types and the Media resource. Conversion from the original data schema to FHIR and update of the application involved a simple Python script for data extraction and transformation followed by import. The processes took minimal effort and time due to the availability of the MediaWiki API.

The use of open-source tools for our DMA platform offers several advantages. The MediaWiki codebase is continually developed and improved by over 500 developers [8]. The open-source community developed the core software and its various extensions, following software engineering best practices with unit/integration testing, version control, and regular releases. It is no surprise that such a large group of engineers is required to support the

engine that powers Wikipedia, one of the world's largest websites. Moreover, MediaWiki is also utilized by many other companies, organizations, and smaller groups to run their public or private wikis. Even a major cancer center has reported the use of a wiki to support part of its operation [20]. It would be difficult to support a large number of individual contributors like that of MediaWiki's for any other software, whether in a company or academic institution setting. As a small informatics team, we benefit from these community developers' contributions with no fees associated with the GPLv2 license [8]. We also leverage other open-source tools for our software infrastructure, including MariaDB for the SQL database and Docker/Kubernetes for application development and deployment. The use of Docker/Kubernetes, in particular, has significantly improved the efficiency of our software development and operation, offering the ease of installation, version control, hardware-agnostic application mobility, and cloud readiness. We have taken advantage of Docker/Kubernetes's flexibility for facile software backups, upgrades, and disaster recovery.

There are, however, some challenges in building with open-source software as opposed to employing a vendor solution. Resources, including a team of engineers, are required to develop and implement the tools. Our department has a platform engineer who has provisioned and maintained our Kubernetes cluster, Docker registry, and networking. Although our DMA application only requires a small server to run, we invested in a cluster of servers to power our Kubernetes system to support the DMA and many other applications. In our use case, the open-source tools are not supported by a vendor; therefore, any issues or bugs associated with the application require our team or the community to troubleshoot and fix. The deployment of the DMA did encounter several issues in the past year while in production use. We have had to address some bugs associated with the extensions that we use in the system, including upload errors with concurrent files. During the initial implementation of our Kubernetes cluster, we encountered a network access problem isolated to one of our laboratories, requiring our platform engineer to work with our central information technology team to correct the issue. A local network outage that affected the entire neighborhood created several days of downtime. Finally, software/hardware incompatibility with our production servers also created days of downtime while migrating the application to alternate servers. We conclude that most of these issues related to network, software, and hardware could have been avoided or perhaps more easily addressed in a more controlled cloud environment. While a vendor may be contractually obligated to address system issues like these, given our experience with past hardware and software problems, this does not necessarily mean that matters will

be resolved promptly or with the required level of customization. Our ability to develop quickly and continuously customize our system to fit our current and future needs outweigh the benefits of using a vendor solution, even with such a small team as ours.

The DMA is one of nine clinical applications that are deployed on a common platform maintained by a departmental informatics team. We estimate that the equivalent of one full-time engineer was required to develop the application over 6 months. In the first year of operation, about 25% of an engineer was needed to further improve the software, address issues, and provide maintenance/infrastructure support. Unless a new feature is being developed or server problems arise, the application after the first year has been issue free with little hands-on intervention and therefore requiring much less effort than 25% of an engineer. Day-to-day access control and assistance with image requests, including the deletion of inadvertently uploaded files, are handled by a system administrator. A web-based visualization dashboard implemented after the launch of our DMA is used to monitor system utilization in real-time.

After a year in clinical production, our department has already reaped the benefits of a web-based platform to manage our gross pathology images. Despite some initial skepticism and reservations about retiring the legacy SND-based system, users across the board, from technologists and PAs to residents and attendings have embraced the new application with very few complaints. They appreciate having quicker access to content for their daily clinical work. Leadership in the department has also valued distributing the image upload labor among pathologist assistants and residents, which has led to the relief of one full-time employee who used to manage image file naming, transfers, and quality control. Beyond clinical utility, the success of our laboratory-developed gross images DMA benefits the department in other ways. Prior to implementation of the DMA, residents preparing for case presentations searched the internet for publicly available gross images. Residents now search our internal DMA to acquire these images as part of the differential diagnoses for their presentations. Several research projects requiring gross images have also leveraged the convenience of the image search capability. We anticipate that both the education and research potential of our system will increase greatly as we expand it to accommodate other image types, and when we integrate additional structured data with the images (e.g., final diagnoses, clinical observations). Our DMA was built with the common FHIR standard, which sets the foundation to adopt other standards as part of a comprehensive solution, including using DICOM to store and retrieve pathology images of all types with associated case and patient metadata.

In conclusion, we have demonstrated the design and implementation of an AP digital media archive solution in

an academic pathology department setting. We show the advantages of using open-source tools to develop a web application based on FHIR structured data for clinical, education, and research use. Such a system provides us the much-needed efficiency and scalability for our current operation and prepares us for integration with other applications and pathology informatics initiatives in the future.

Data availability

The datasets used and analyzed during the current study are not available.

Compliance with ethical standards

Conflict of interest The authors declare no competing interests.

Ethical approval This paper describes a quality improvement project and did not require ethical approval.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

References

1. Rampy BA, Glassy EF. Pathology gross photography: the beginning of digital pathology. Surg Pathol Clin. 2015;8:195–211.
2. Horn CL, DeKoning L, Klonowski P, Naugler C. Current usage and future trends in gross digital photography in Canada. BMC Med Educ. 2014;14:11.
3. Riley RS, Ben-Ezra JM, Massey D, Slyter RL, Romagnoli G. Digital photography: a primer for pathologists. J Clin Lab Anal. 2004;18:91–128.
4. Belanger AJ, Lopes AE, Sinard JH. Implementation of a practical digital imaging system for routine gross photography in an autopsy environment. Arch Pathol Lab Med. 2000;124:160–5.
5. Marchevsky AM, Dulbandzhyan R, Seely K, Carey S, Duncan RG. Storage and distribution of pathology digital images using integrated web-based viewing systems. Arch Pathol Lab Med. 2002;126:533–9.
6. Park RW, Eom JH, Byun HY, Park P, Lee KB, Joo HJ. Automation of gross photography using a remote-controlled digital camera system. Arch Pathol Lab Med. 2003;127:726–31.
7. Sinard JH, Mattie ME. Overcoming the limitations of integrated clinical digital imaging solutions. Arch Pathol Lab Med. 2005; 129:1118–26.
8. GitHub: wikimedia/mediawiki, [Internet], [cited 13 February 2021]. Available from https://github.com/wikimedia/mediawiki.
9. HL7. Resource media—content, [Internet], [cited 13 February 2021]. Available from https://www.hl7.org/fhir/media.html.
10. Barut C, Ertilav H. Guidelines for standard photography in gross and clinical anatomy. Anat Sci Educ. 2011;4:348–56.
11. O'Brien MJ, Sotnikov AV. Digital imaging in anatomic pathology. Am J Clin Pathol. 1996;106:S25–32.
12. Amin M, Sharma G, Parwani AV, Anderson R, Anderson BJ, Piccoli A, et al. Integration of digital gross pathology images for enterprise-wide access. J Pathol Inform. 2012;3:10.
13. Hong MK, Yao HHI, Pedersen JS, Peters JS, Costello AJ, Murphy DG, Hovens CM, et al. Error rates in a clinical data repository: lessons from the transition to electronic data transfer-a descriptive study. BMJ Open. 2013;3:e002406.
14. Vecellio E, Malley MW, Toouli G, Georgiou A, Westbrook JI. Data quality associated with handwritten laboratory test requests: classification and frequency of data-entry errors for outpatient serology tests. Health Inf Manag. 2015;44:7–12.
15. Bauer JC, John E, Wood CL, Plass D, Richardson D. Data entry automation improves cost, quality, performance, and job satisfaction in a hospital nursing unit. J Nurs Adm. 2020;50:34–9.
16. MediaWiki, [Internet], [cited 13 February 2021]. Available from https://en.wikipedia.org/wiki/MediaWiki.
17. The top 500 sites on the web, [Internet], [cited 13 February 2021]. Available from https://www.alexa.com/topsites.
18. HL7 FHIR Overview, [Internet], [cited 13 February 2021]. Available from https://www.hl7.org/fhir/overview.html.
19. Clunie DA, Dennison DK, Cram D, Persons KR, Bronkalla MD, Primo HR. Technical challenges of enterprise imaging: HIMSS-SIIM collaborative white paper. J Digit Imaging. 2016;29:583–614.
20. Koren Y. Between the brackets: episode 35: Alex Tanchoco. https://betweenthebrackets.libsyn.com/episode-35-alex-tanchoco.