

Introduction to the DICOM standard for digital pathology and its importance for workflow efficiency

Digital pathology is proving to provide large benefits for healthcare providers, but growing adoption is now presenting these early adopters with new challenges. The ability of slide scanners to communicate with surrounding IT systems regardless of vendor, or the other way around, for IT systems to communicate with scanners regardless of vendor, is one of these major challenges. Historically, scanners have in general produced their own specific file types, which are only viewable with the scanner vendor's proprietary viewer, or in the best case scenario, with selected partners' viewers to which they have opened up their file format. A standardized file format for digital pathology images is key for achieving highly efficient workflows.

As the adoption of digital pathology – and thus the digital workload – is growing rapidly, it is becoming increasingly important not to be tied to a specific workstation or scanner. An efficient pathology workflow is dependent on interoperability between different scanners and IT systems. This is therefore one of the major drivers behind the adoption of the DICOM standard in pathology.

This white paper gives an introduction to the DICOM standard and its role in facilitating an efficient pathology workflow, highlighting the standard's main advantages and limitations.

Why the DICOM standard was created

In the early days of digital radiology, imaging devices from different vendors produced proprietary file formats only compatible with their own PACS¹ and viewers, which made it practically impossible for radiology departments to replace their PACS with one from a different vendor without also exchanging all their modalities. This incompatibility was one of the main drivers behind the implementation of a standard called DICOM². Compliance with the DICOM standard meant that modalities produced images on a standardized format and communicated with stan-

dard protocols. The DICOM standard has also been successfully implemented in other disciplines using digital medical imaging, and is now on its way to be adopted in digital pathology.

Advantages

Integration with existing IT systems

Incompatibility issues and the lack of a singular method to acquire and store pathology's large images has been a major stumbling block in the acceptance of digital pathology. The DICOM Standard Committee Working Group 26, consisting of scanner vendors, PACS vendors and pathologists, has put in a tremendous effort to standardize storage methods so that they are more in line with currently available PACS in most hospitals for storage of radiology images³. The incorporation of modern digital pathology into the DICOM standard was finalized in 2010. It was described in Supplement 145 and was praised when released since it allowed hospitals to integrate digital pathology into their existing IT systems without adding too many costs⁴. The DICOM format also makes it easier for pathology images to be stored together with images from other specialties in the same archive, which becomes important as more and more healthcare providers implement so-called VNAs⁵.

Preventing vendor lock-in

The main advantage with DICOM is that it is a standard. Using a standard for file format and communication allows the pathology department to connect scanners and PACS from different vendors and replace them without incompatibility problems.

“If the lab has produced a large archive of digital pathology images and these are in DICOM, the lab can replace the PACS and scanners and still review the images produced by the old scanners stored in the previous PACS.”

Swedish Pathologist

Cross-discipline collaboration

Adopting DICOM in digital pathology would also lower the barriers for collaboration between pathology, radiology, surgery and radiation therapy since images from these disciplines could be viewed and managed in each other's systems. This would unlock the true benefits of integrated diagnostics and allow, for example, for enhanced collaboration in multidisciplinary conferences, as well as making it possible to detect discrepancies in findings and provide a more holistic patient overview.

High performance review and image handling

Because of how information is stored in DICOM files, the DICOM format will in most cases show higher performance in displaying and viewing images.

Supplement 145 recommends a simple approach to storing the large multi-resolution multi-planar images in pathology and mapping sub-regions from each layer into a DICOM series. These sub-regions are referred to as tiles, in which images are stored in squares (or

rectangular) tiles, which are then stored in two-dimensional arrays.

This approach is basically like cutting a large picture into smaller squares or rectangular fragments and then storing them together in a box where multiple resolutions are stored in the same way. The highest resolution has the most data and will form the broadest row, and each row of less resolution will stack up on top of rows containing more resolution, forming a sort of pyramid. Each layer can then be retrieved and put together, to either form part of an image or the entire picture. This makes it relatively easy to randomly access any sub-region of the image without loading large amounts of data, as well as to build the images on pre-computed tiles in multiple resolutions. All in all, this allows for faster viewing of the image.

Some of the proprietary file formats also utilize pre-computed tiles in multiple resolutions, but viewing performance tests conducted by Sectra have shown that it is often possible to achieve better performance with DICOM images than with proprietary file formats. One of these tests is shown in graph 1, where viewing performance has been measured for both solid state drive (expensive) and standard hard disk drive (cheap). Viewing performance is measured as the response time on the x-axis and the number of tiles fetched on the y-axis; hence, a high performance view of an image is indicated by having the majority of requests within 20 ms. From the graph it is clear that the difference is small between reading a DICOM file and a proprietary file that has been optimized for reading. However, many proprietary file formats are optimized for scanning (writing), which gives a significant worse reading performance on the cheaper standard hard disk drive.

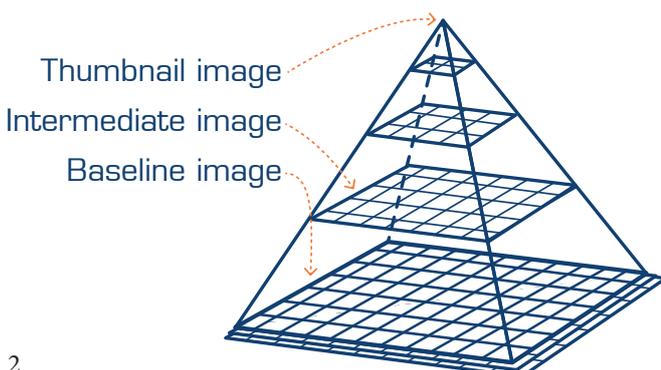
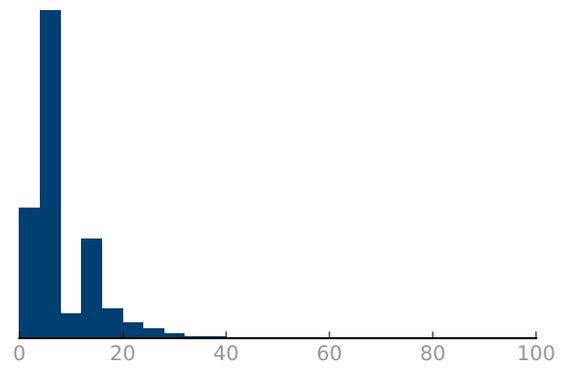
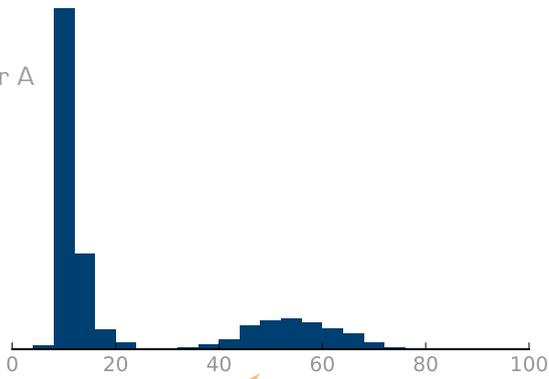


Figure 1: The highest resolutions captured by an image scanner will sit at the base of the pyramid, while the apex corresponds to the zoom or low power image⁶.

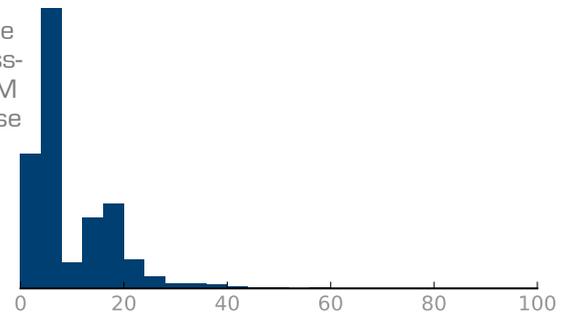
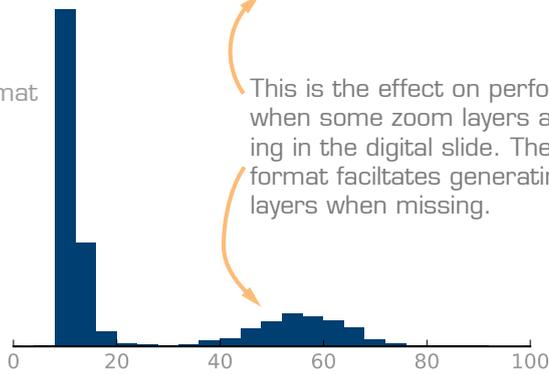
Standard hard disk drive

Solid-state drive

DICOM
Converted from vendor A

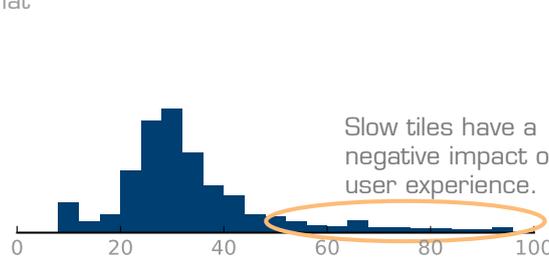


Vendor A
Reading optimized format

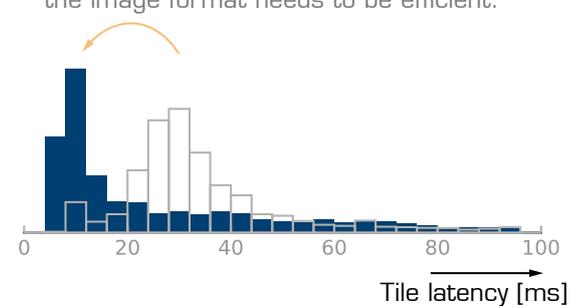


This is the effect on performance when some zoom layers are missing in the digital slide. The DICOM format facilitates generating these layers when missing.

Vendor B
Writing optimized format



A faster disk shifts the tile loading times to the left. In order to be able to use cheaper and slower disk, the image format needs to be efficient.



Slow tiles have a negative impact on user experience.

Graph 1: One of the viewing performance tests conducted by Sectra.

Smarter scanning and review

Scanners do not always capture all areas of the slide at the highest resolution. These are called sparse images and can be handled by the DICOM standard. The DICOM standard allows the image to exclude pixel data for uncaptured areas, resulting in a smaller image, which is beneficial for the performance in viewing and reduces scanning time.

More “informed” system with the DICOM header

The DICOM header includes tags where image-specific information and patient information is saved. Any system reading the image will have the possibility to access, process and display that data. This can, for example, be utilized when applying image analysis algorithms where information in the header can steer pre-processing of the image after specific protocols (e.g. prostate) or give the user access to different tools depending on organ type (e.g. number of cells counted in Ki67).

The DICOM header information also enables the image to be connected to the right referral and patient in the PACS. This reduces the reliance on the right information being contained in the barcode on the glass and on that the barcode being scanned correctly.

In a migration scenario, where the images are moved to a different PACS or VNA, the DICOM header also facilitates faster migration with better data quality and less risk of errors.

Utilizing existing DICOM-based functionality from radiology PACS

There is a significant amount of DICOM-based functionality in a radiology PACS that pathology can utilize as a foundation for future development if adopting the DICOM standard. However, DICOM was originally designed for storing and forwarding multiple modest-size images, and the same is true of the DICOM-based functionality for sharing radiology images. Due the large image size in pathology, certain PACS functionalities need to be developed further and complemented with new technology to work in pathology. For example, image streaming technology will be crucial for efficient remote reading of digital pathology images. Remote pathology readers will experience the best performance by streaming digi-

tal pathology images directly from the PACS, which only requires 2-3%⁷ of the image to be transmitted.

Limitations

Currently slower scanning

In general, scanners has been optimized for speed in producing their proprietary files (writing). If scanners are not redesigned when adopting the DICOM standard – for example, when it comes to integration with the LIMS – an additional step needs to be added to the process, which might prolong the scanning time.

Requires additional indexing to provide high viewing performance

In order for a viewer to know where in a DICOM instance to look for a particular part of the image, all per-frame information in the DICOM header needs to be parsed. This could, if implemented straight off, potentially lead to long waiting times before an image is viewable.

However, this can be mitigated by performing necessary parsing before the image needs to be viewed, for instance, when the image is imported. The frames in the DICOM header can be indexed in an efficient manner so that access to a particular part of the image is near instantaneous. The pre-parsing step needs to create an index that makes it possible to quickly look up the file offset to the relevant DICOM frame, given the coordinates of a tile. Something similar to the TIFF format's indexing of tiles using the tile offsets and tile length tags would work. This is not part of the DICOM standard though, and needs to be handled by applications separately.

Populating the DICOM header may be difficult

The DICOM header contains some tags that are mandatory and some that are optional. Scanner vendors today struggle to populate all mandatory tags into the DICOM header. In radiology, the modalities are normally fed by a worklist from the information system over a protocol called DICOM Modality Worklist. In a similar way, one could argue that the pathology scanners could be fed by the LIS to steer the scanning and populate tags in the DICOM header. However, this has been shown to be difficult since few LIS solutions have implemented a modality worklist interface yet. Scanning also differs from pro-

ducing radiology images in that it is batch oriented and there is no moment where the scanner operator can supply details about a current patient.

Some pathologists also see it as a disadvantage that the DICOM header contains patient information, since pathologists are used to the anonymous glass slides.

In addition, the DICOM header makes the files on average 4-5% larger than the average proprietary file based on a .tiff-format.

The standard is open for interpretation

The DICOM standard is very flexible and permissive. This can be an advantage as it allows for flexibility among vendors in order to optimize each specific function or system, but it also incurs a risk of diverging file formats, which might cause problems in terms of interoperability. In the future, it is clear that the standard will require best practices to avoid conflicts in interpretation. Examples, and Sectra's suggestions for best practices, include:

» Performance

- Choose a good tile size, balancing the need to fill a high-resolution monitor with as few network round-trips as possible while still not having to fetch image data that won't be displayed. 512px x 512px is probably a good tile size with 4MP or 8MP monitors.

- Let the scanner generate enough pyramid levels in order to avoid the need for on-the-fly down sampling of large amounts of high-resolution data.

» Simplicity

- Use quadratic tiles.
- Use the same tile size for all zoom levels.
- Arrange tiles in manner that makes it possible to easily find the location from which to read a particular part of the image. Avoid splitting a pyramid level spatially, unless necessary due to usage of concatenated DICOM instances.

» Viewing performance

- When making choices during implementation of DICOM support in scanners, we recommend that scanner vendors consider viewer performance. A slide might be scanned once, but viewed multiple times with high demands on performance. To increase interoperability and the likelihood of viewers and PACS supporting the scanner images, vendors should try favoring simplicity in their interpretation and implementation.



“ For DICOM in digital pathology to be successfully implemented, we need to learn from how similar issues were solved in radiology, listen to vendors and user feedback, and learn from early implementations. ”

Conclusion

Although many challenges remain, DICOM is being implemented in digital pathology and there are no doubts that it will be the standard. Scanner and PACS vendors are currently taking a step-by-step approach towards full adoption of the DICOM standard.

The *first step* involves format compliance. Scanners need to produce the DICOM images, and PACS vendors need to be able to import and view these files. The *second step* involves supporting the DICOM standard in terms of communication. The *third step* is to enable the already existing DICOM-based functionality in other IT systems, such as PACS and VNAs, to handle some of the larger pathology images.

For DICOM in digital pathology to be successfully implemented, we need to learn from how similar issues were solved in radiology, listen to vendors and user feedback, and learn from early implementations. For the DICOM standard to be successful in pathology, we have concluded the following:

- » DICOM images can in general be viewed with higher performance than proprietary files, which is a strong enough argument to start using the standard. The standard is already good enough to begin adoption in order to ensure compatibility between systems and ensure future-proof storage of digital pathology images.
- » The standard must be implemented properly since, when applied in the right way, DICOM will provide a fast image format that support systems with a high workload and high demands on viewing.
- » Since the standard is very flexible and permissive, the industry actors should develop a best practice on how to implement DICOM to guide all new development towards a harmonized approach.

Reference List

1. *Picture Archiving and Communication System* – IT system used for review, archive and distribution of radiology, pathology and other images in healthcare
2. *Digital Imaging and Communications in Medicine*
3. <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3097525/>
4. <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3097525/>
5. *Vendor Neutral Archive* – an archive which could archive images from several different specialties
6. <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3097525/figure/F1/>
7. Measurement made on Sectra installation