Publication

Privacy risks of whole-slide image sharing in digital pathology

May 2023

In their recent publication, researchers from BBMRI.at partner Med Uni Graz propose a model for the assessment of privacy risks associated with the sharing of whole-slide images in the growing field of digital and AI-assisted pathology.

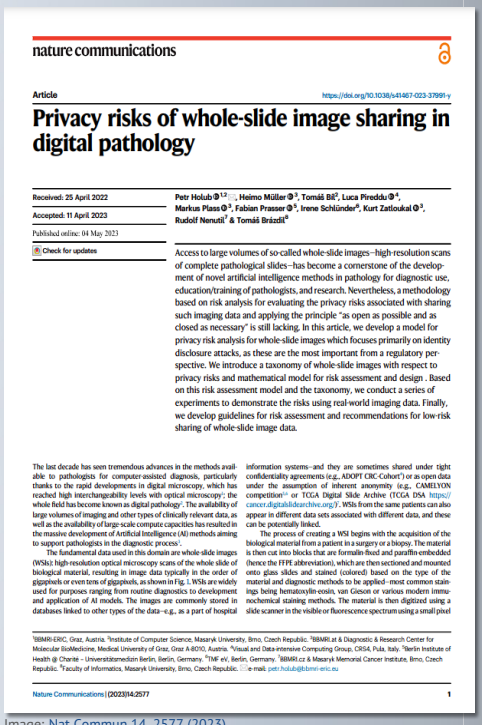
In their paper regarding security risks in digital pathology, experts from BBMRI.at partner Med Uni Graz introduce a hierarchical taxonomy of whole-slide images based on the potential of these images to be linked to each other. In addition, they provide a thorough mathematical model aimed at assessing the risk of security threats associated with the sharing of whole-slide images and related data, especially regarding identity disclosure attacks. Furthermore, these risks are demonstrated using real-life data sets, evaluating the newly established risk assessment model in a series of experiments.

The authors conclude their findings by setting up detailed guidelines for the release of whole-slide images as anonymous or personal data and state their recommendations for the minimization of privacy risks in the context of sharing whole-slide images for digital pathology applications.

Read publication:

Holub, P.; Müller, H.; Bíl, T.; Pireddu, L.; Plass, M.; Prasser, F.; Schlünder, I; Zatloukal, K.; Neunutil, R.; Brázdil, T. Privacy risks of whole-slide image sharing in digital pathology. Nat Commun 14, 2577 (2023). https://doi.org/10.1038/s41467-023-37991-y

More publications from BBMRI.at partners>



nature communications

Article

https://doi.org/10.1038/s41467-023-37991-y

Privacy risks of whole-slide image sharing in digital pathology

Received: 25 April 2022

Accepted: 11 April 2023

Published online: 04 May 2023

Check for updates

Petr Holub, Heimo Müller, Tomáš Bíl, Luca Pireddu, Markus Plass, Fabian Prasser, Irene Schlünder, Kurt Zatloukal, Rudolf Nenutil & Tomáš Brázdil

Access to large volumes of so-called whole-slide images—high-resolution scans of complete pathological slides—has become a cornerstone of the development of novel artificial intelligence methods in pathology for diagnostic use, education/training of pathologists, and research. Nevertheless, a methodology based on risk analysis for evaluating the privacy risks associated with sharing such imaging data and applying the principle "as open as possible and as closed as necessary" is still lacking. In this article, we develop a model for privacy risk analysis for whole-slide images which focuses primarily on identity disclosure attacks, as these are the most important from a regulatory perspective. We introduce a taxonomy of whole-slide images with respect to privacy risks and mathematical model for risk assessment and design . Based on this risk assessment model and the taxonomy, we conduct a series of experiments to demonstrate the risks using real-world imaging data. Finally, we develop guidelines for risk assessment and recommendations for low-risk sharing of whole-slide image data.

The last decade has seen tremendous advances in the methods available to pathologists for computer-assisted diagnosis, particularly thanks to the rapid developments in digital microscopy, which has reached high interchangeability levels with optical microscopy[1]; the whole field has become known as digital pathology[2]. The availability of large volumes of imaging and other types of clinically relevant data, as well as the availability of large-scale compute capacities has resulted in the massive development of Artificial Intelligence (AI) methods aiming to support pathologists in the diagnostic process[3].

The fundamental data used in this domain are whole-slide images (WSIs): high-resolution optical microscopy scans of the whole slide of biological material, resulting in image data typically in the order of gigapixels or even tens of gigapixels, as shown in Fig. 1. WSIs are widely used for purposes ranging from routine diagnostics to development and application of AI models. The images are commonly stored in databases linked to other types of the data—e.g., as a part of hospital information systems—and they are sometimes shared under tight confidentiality agreements (e.g., ADOPT CRC-Cohort[4]) or as open data under the assumption of inherent anonymity (e.g., CAMELYON competition[5,6] or TCGA Digital Slide Archive (TCGA DSA https://cancer.digitalslidearchive.org/)[7]. WSIs from the same patients can also appear in different data sets associated with different data, and these can be potentially linked.

The process of creating a WSI begins with the acquisition of the biological material from a patient in a surgery or a biopsy. The material is then cut into blocks that are formalin-fixed and paraffin-embedded (hence the FFPE abbreviation), which are then sectioned and mounted onto glass slides and stained (colored) based on the type of the material and diagnostic methods to be applied—most common stainings being hematoxylin-eosin, van Gieson or various modern immunochemical staining methods. The material is then digitized using a slide scanner in the visible or fluorescence spectrum using a small pixel

[1]BBMRI-ERIC, Graz, Austria. [2]Institute of Computer Science, Masaryk University, Brno, Czech Republic. [3]BBMRI.at & Diagnostic & Research Center for Molecular BioMedicine, Medical University of Graz, Graz A-8010, Austria. [4]Visual and Data-intensive Computing Group, CRS4, Pula, Italy. [5]Berlin Institute of Health @ Charité – Universitätsmedizin Berlin, Berlin, Germany. [6]TMF eV, Berlin, Germany. [7]BBMRI.cz & Masaryk Memorial Cancer Institute, Brno, Czech Republic. e-mail: petr.holub@bbmri-eric.eu

Image: Nat Commun 14, 2577 (2023)

Bundesministerium Bildung, Wissenschaft und Forschung

Funded by
GZ 10.470/0016-II/3/2013 (2013-2018)
BMBWF-10.470/0010-V/3c/2018 (2018-2023)