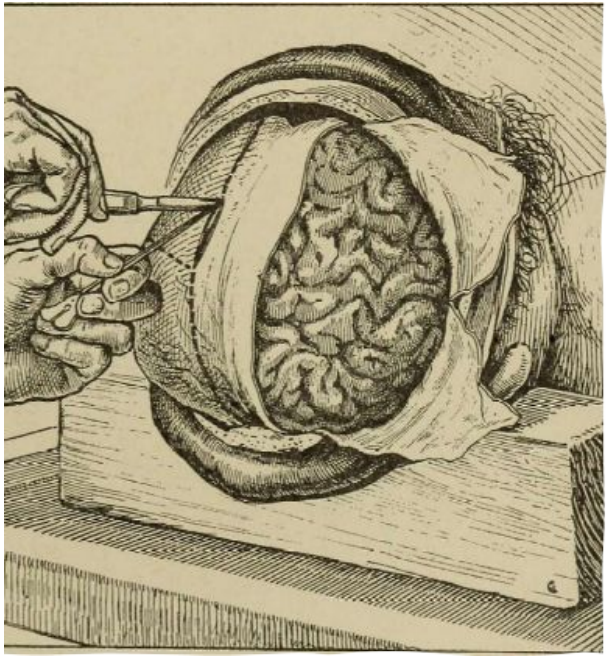
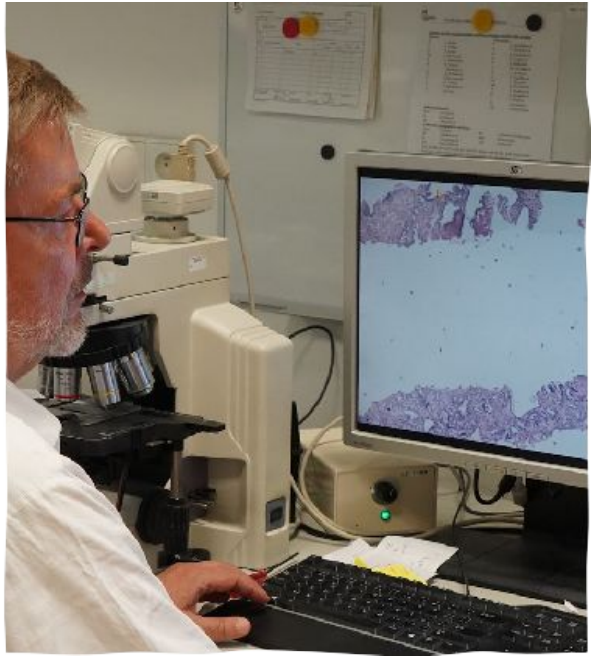
Artificial Intelligence in Digital Pathology

From Concept to Clinical Practice

Tomáš Brázdil

Evolution of Pathology





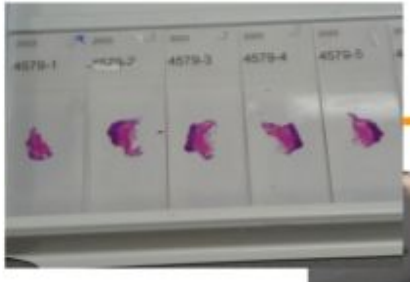
… from cutting dead bodies to advanced diagnostics

What pathologists do?

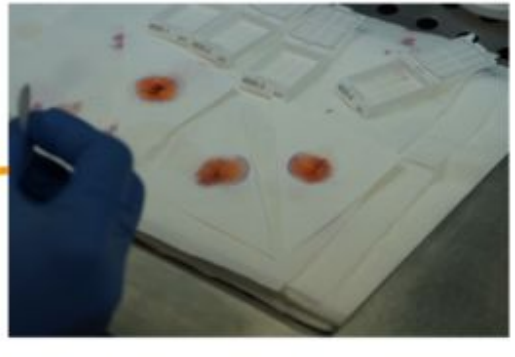
Scanner

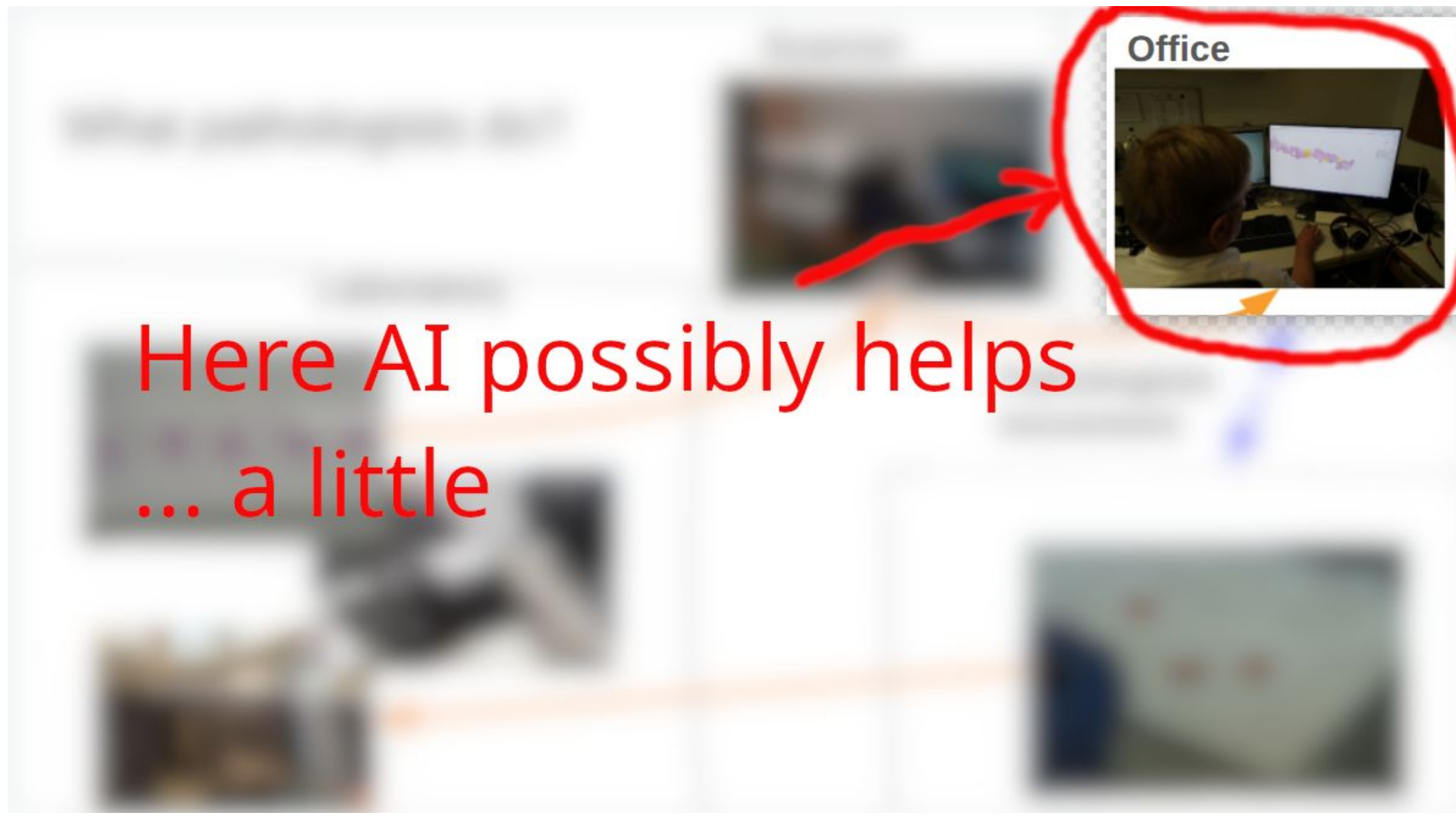
Office

Laboratory

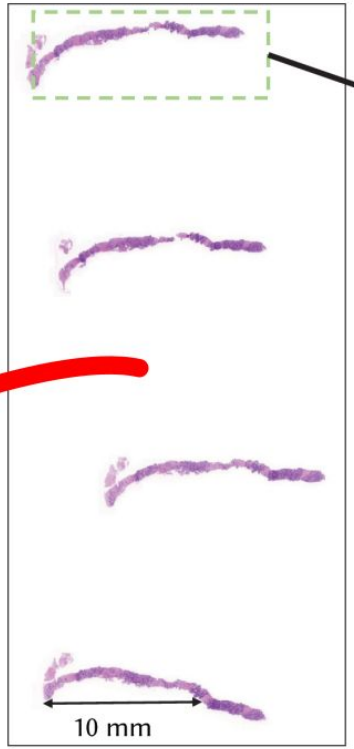


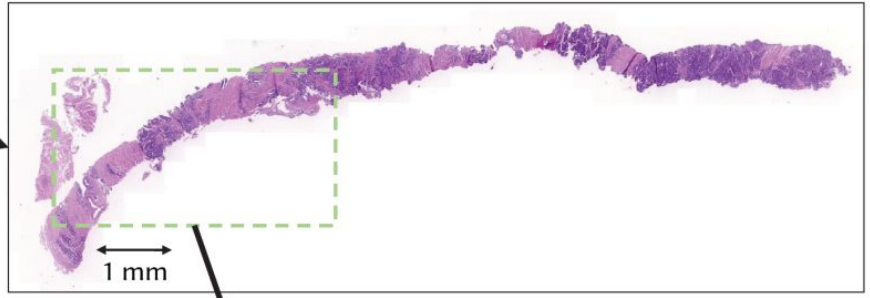
Pathologist's movement

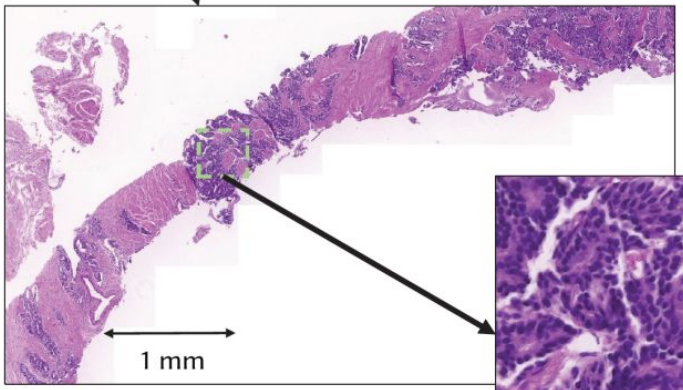


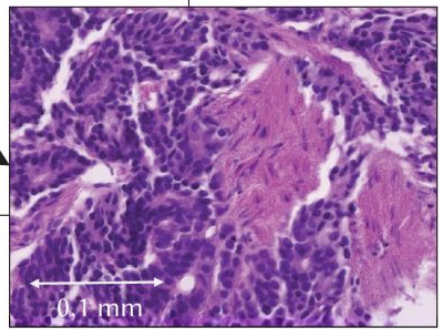

Office
Here AI possibly helps
… a little

What a pathologist sees?

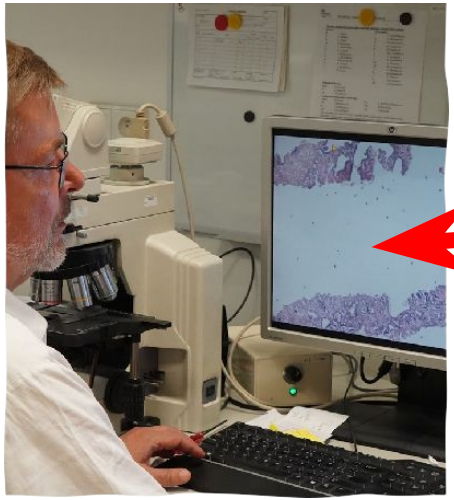


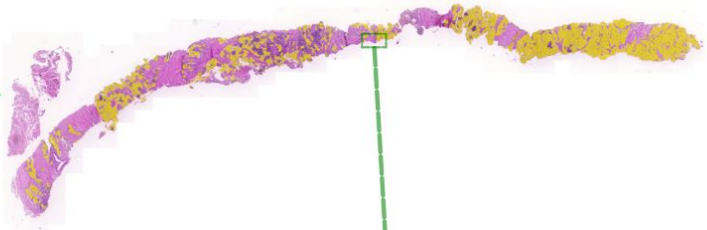


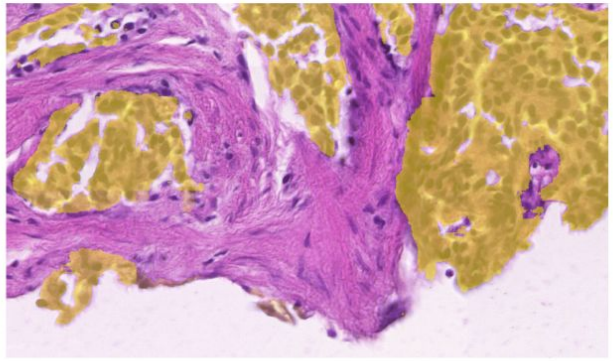




AI power!!



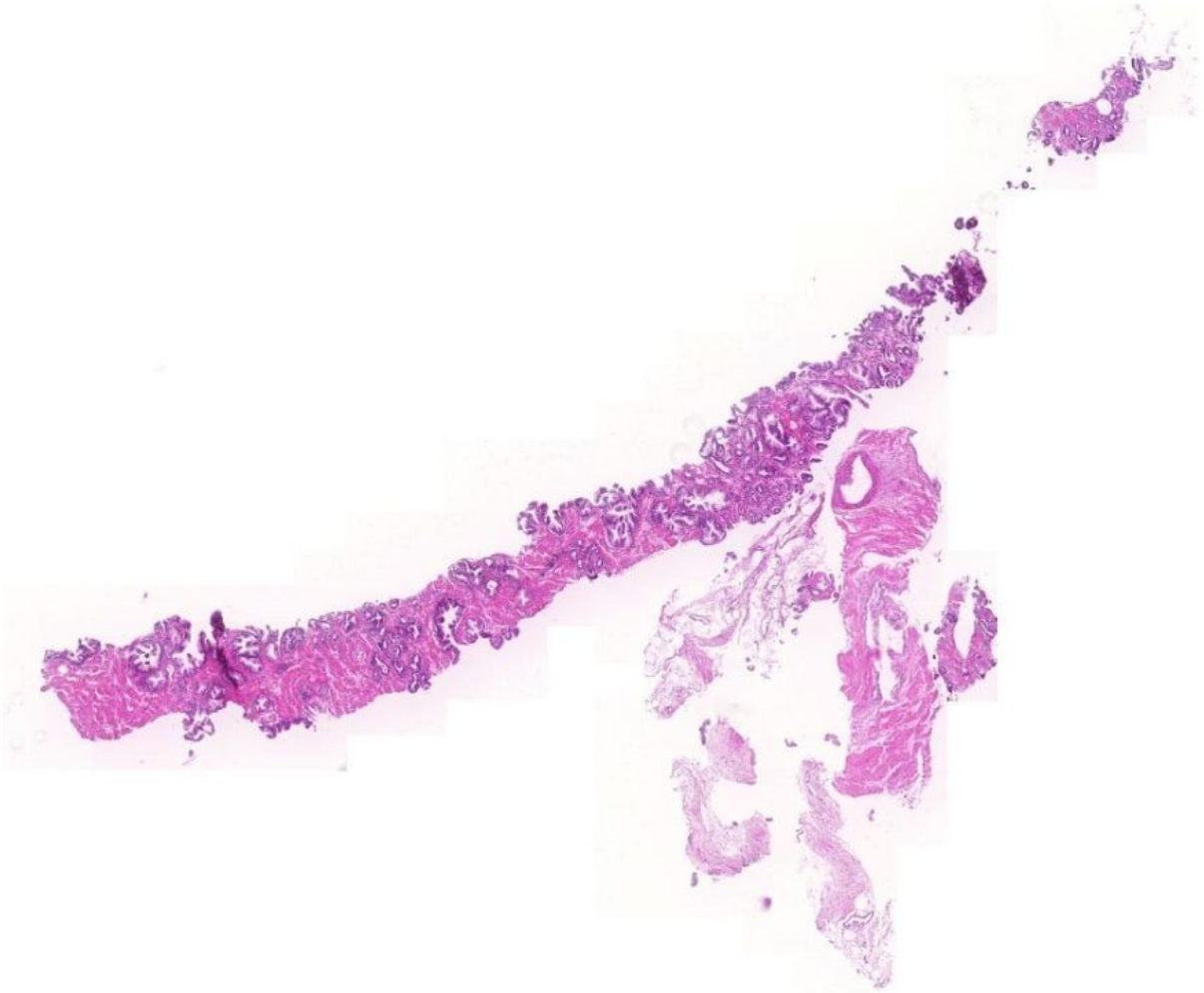




Data

Where is the cancer?

Beware, huge image!!

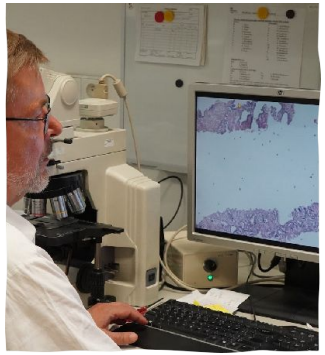


Data

Cancer is approx. here

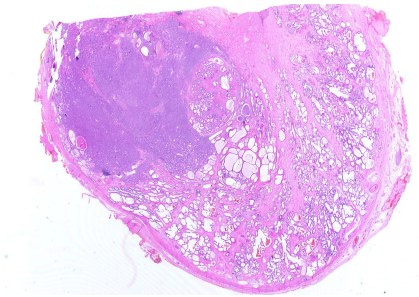
Train AI to find the cancer!

Digital pathology



Data characteristics

- *Image data*: **Whole-slide images**
  - gigapixel images of tissue
  - several images per case
- *Textual data*: **Medical reports**
  - unstructured, differ in style, language etc.
  - medical jargon
- *Tabular data*: **Clinical, genomic, etc.**
  - Not standardized across institutions



Microscopic Description: Histologic examination reveals a soft tissue specimen consisting of keratinized stratified squamous epithelium and underlying connective tissue. The epithelium exhibits a thin surface layer of parakeratin subjacent to which the spinous cell layer varies in thickness. The basal cell layer is disorganized or nonexistent throughout much of the specimen. The underlying connective tissue is comprised of delicate to dense bundles of intertwining collagen fibers interspersed by varying numbers of fibrocytes and small blood vessels. Prominent within this framework is a band-like infiltrate of lymphocytes present immediately subjacent to the epithelium and focally obscuring the epithelial/connective tissue junction. In some of these areas, the epithelium is artifactually separated from the underlying connective tissue. Superficially, bacterial colonies are adherent to the epithelium.

Patient no.	Harvey–Bradshaw index	Inflammatory Bowel Disease Questionnaire	C-reactive protein (mg/L)	Highest CTE score	Highest SUVmax	Highest SUVmax ratio	Highest simplified endoscopic score	Highest pathology inflammation grade
1	4	125	15.8	7	6.4	4.6	4	4
2	3	156	8.9	0	3.6	1.2	3	2
3	5	167	1.9	5	5.8	3.4	3	3
4	3	196	29	5	7.5	4.0	NA*	4
5	10	181	NA	5	6.6	3.3	NA	4
6	11	138	NA	9	5.9	4.9	NA	4
7	2	188	NA	0	3.4	1.2	1	2

AI training workflow

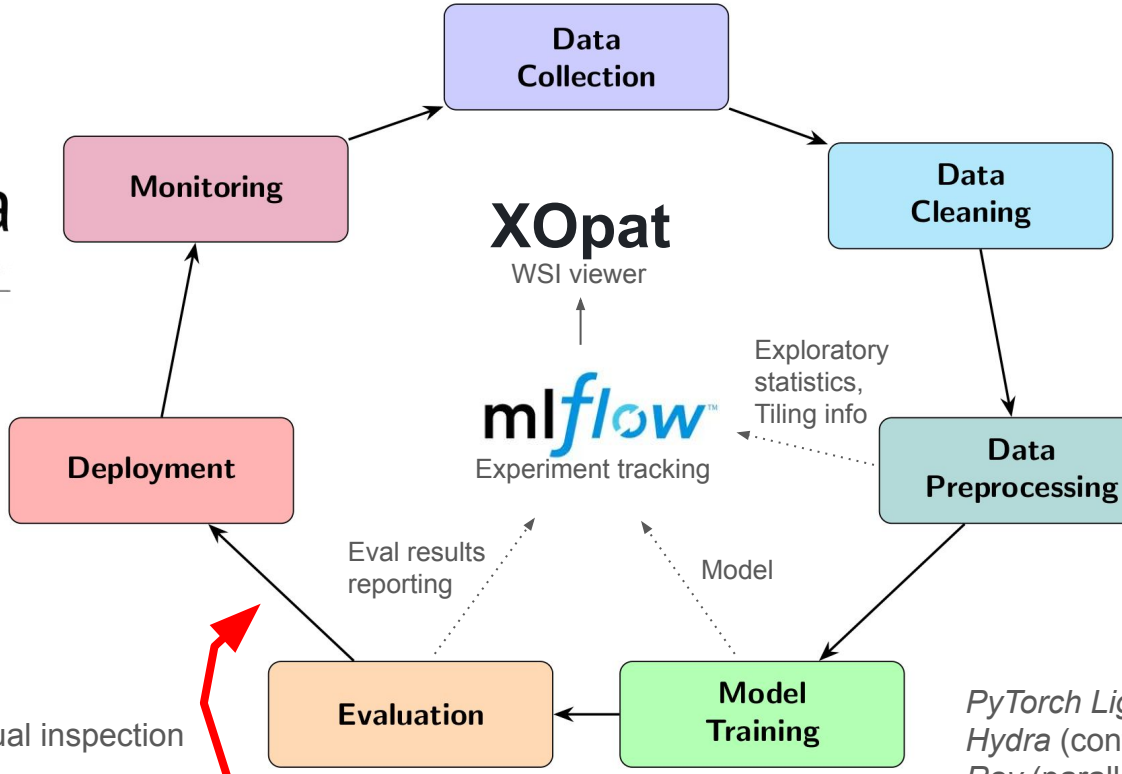
- The **pathologist formulates** the **medical problem**

  E.g. detect prostate cancer in whole-slide images (WSI) of prostate needle biopsies
- The **pathologist selects** the appropriate **cases/slides** for AI training
- The **slides** are **scanned**
  - No standard of WSI format
  - Various scanning parameters
- **Metadata** collected and processed
- The resulting data handed over to computer scientists
  - smaller projects = thousands of WSI (terabytes)
  - large projects = hundreds of thousands/millions of WSI (petabytes)
- Training/evaluation results **evaluated** by the **pathologist**
- Successful models are **deployed** and **monitored**

WSI processing AI

Transferring WSI and metadata from hospitals
*Storing at **Cerit-SC***

empaia
International e.V.
ecosystem for pathology diagnostics with AI assistance



WSI reading (OpenSlide)
WSI format unification
Metadata quality check
WSI quality check
Staining consistency, folds, blurring, etc.

Data exploration
Tiling
Augmentations

PyTorch Lightning
Hydra (config management)
Ray (parallel computing)

TorchMetric
xOpat viewer for visual inspection

“Easy” up to this point

Illustration: Prostate cancer detector

Training

- 785 WSI
  166 patients,
  detailed annotations
- Cut into 7,878,675
  small patches
- Trained a simple
  VGG16 binary model

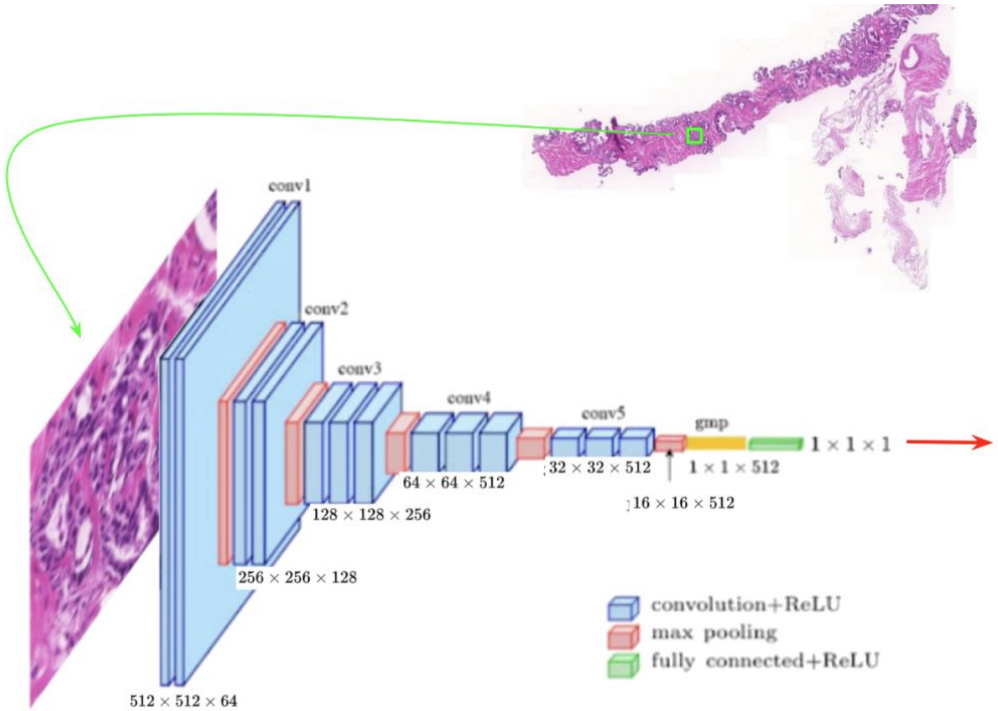


Illustration: Evaluation

- (Almost) perfect results on a test set of 90 WSI
- Visually the model goes for known cancer features

… but is it really useful??

Solved by AI?

Pathologist's view: It does not bother me and sometimes it's helpful

IT expert's view: In 2018

cancer patients with clinicopathological and outcome data available. The results show that deep learning-based outcome prediction with only small tissue areas as input outperforms (hazard ratio 2.3; CI 95% 1.79–3.03; AUC 0.69) visual histological assessment performed by human experts on both TMA spot (HR 1.67; CI 95% 1.28–2.19; AUC 0.58) and whole-slide level

Bychkov et al (2018). Deep learning based tissue analysis predicts outcome in colorectal cancer. *Scientific Reports*, 8

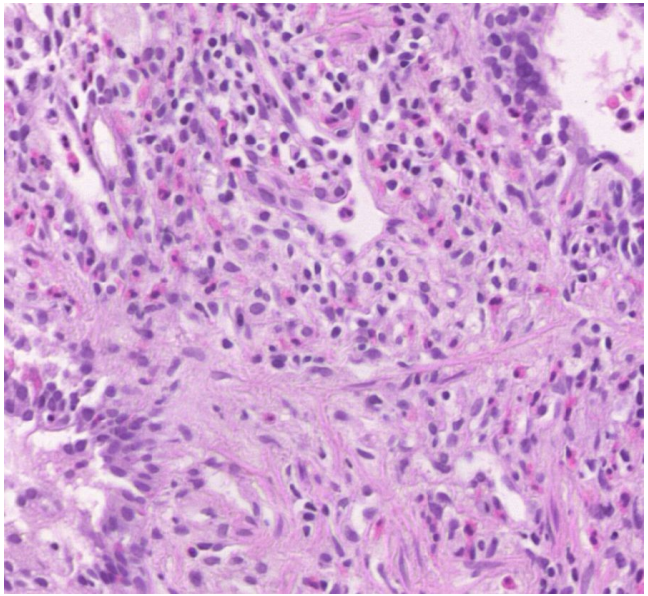
Solved by AI?

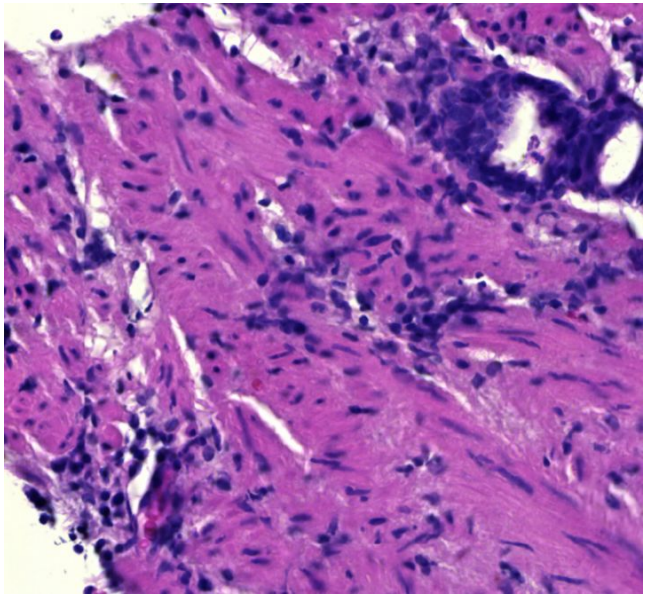
IT expert's view: In 2024 (and still in 2025)

> Researchers have published many promising algorithmic solutions.[11,12] However, the path to wide clinical adoption is difficult. A core problem is a lack of standardization and interoperability for the seamless integration of image analysis methods into diverse image management and laboratory information systems. Commercialization and clinical implementation of pathology AI must overcome additional hurdles,[13,14] namely the transformation of an idea into an AI prototype (which requires data acquisition), a validation process towards market readiness, and certification as a medical product. Finally, reimbursement and billing issues must be solved to generate revenue.

Zerbe et al (2024). Joining forces for pathology diagnostics with AI assistance: The EMPAIA initiative. *Journal of Pathology Informatics*, 15

Let’s go to another hospital …





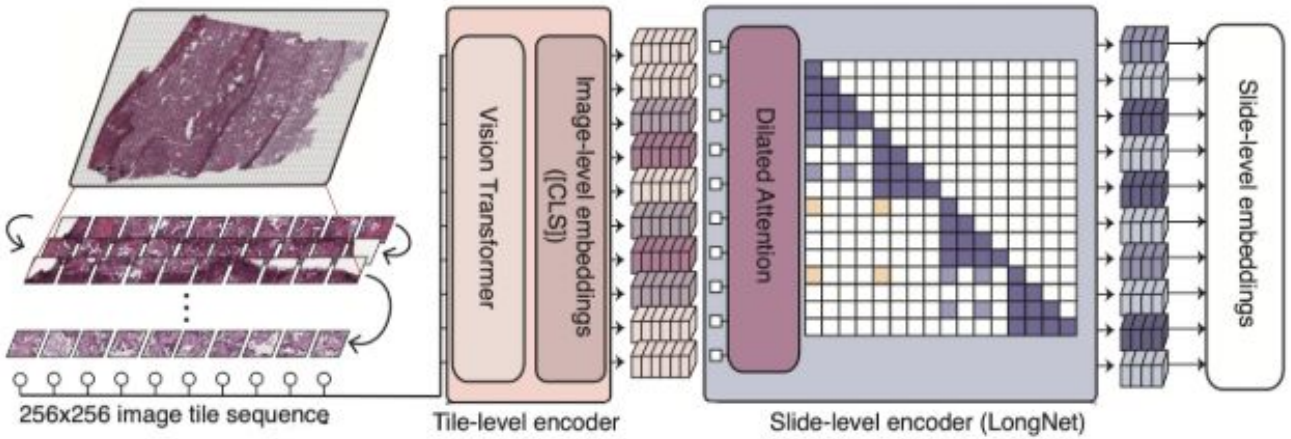
AI trained on the MMCI data (left) fails significantly on FH Brno data (right)

What is the difference?

Modern solution

- Large foundation models
  - Trained on 100Ks/millions of WSI of various stainings, diagnosis etc.
  - Extract generic features relevant to pathology
  - Features fed into smaller models solving particular problems

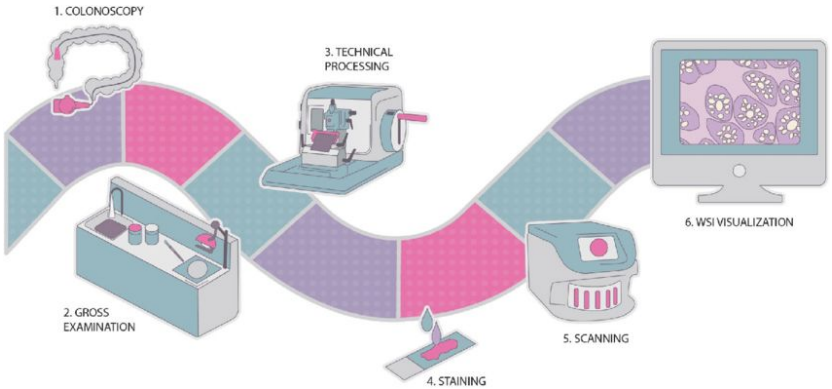
Prov-GigaPath
28 cancer centres
30,000 patients
31 tissue types



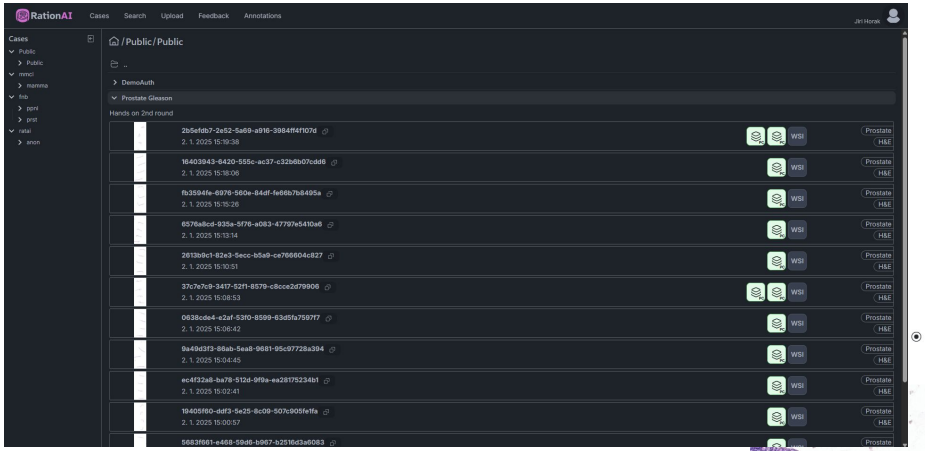
Xu et al, A whole-slide foundation model for digital pathology from real-world data. Nature, 2024

What is so hard about the deployment?

- Technology
  - Different hospital/laboratory information system
  - Storage (each WSI several GB), compute (large networks = large GPUs)
  - Ergonomy
- Management
  - AI workflow integration
  - Pathologist's refusal, AI reliability
  - Competency building
- Certification
  - IVDR -> expensive, difficult, long-winded



Experimental clinical deployment at MMCI

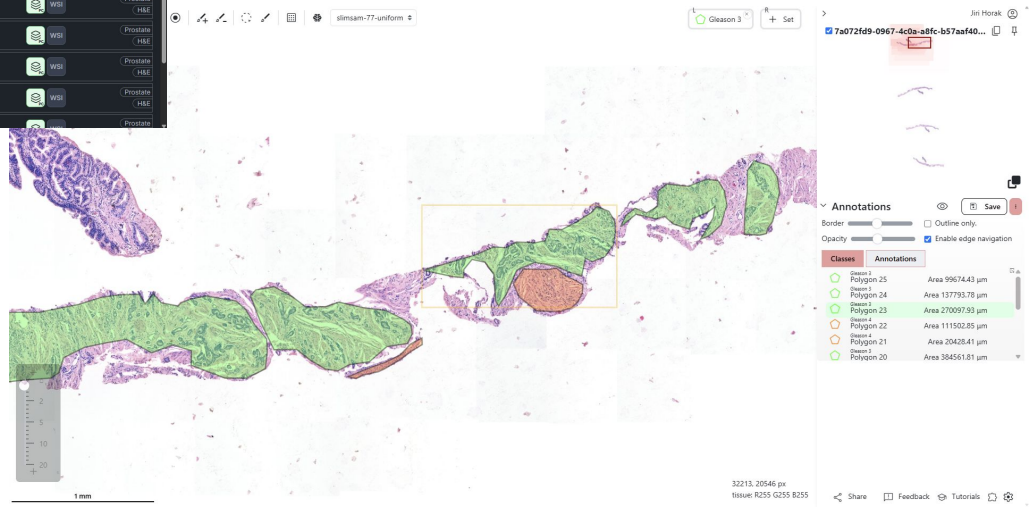


Case browser

- Execution of AI models
- Metadata examination

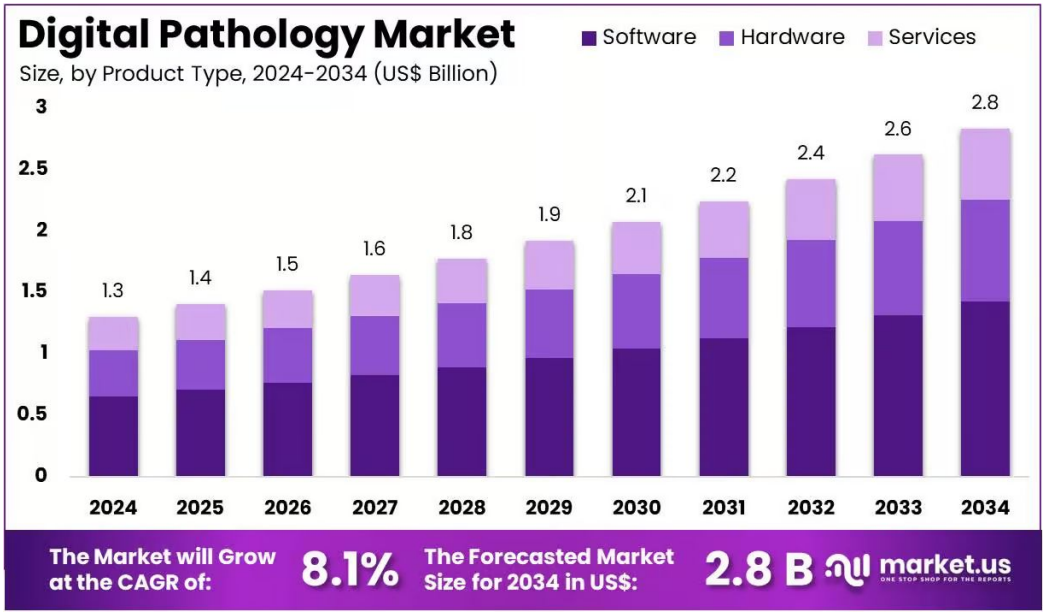
xOpat viewer

- Display scanned samples
- Annotations, model outputs



… and the money?

- Relatively small highly specialized market
- Costly and lengthy certification process
- Vendor dependent scanner software
- Clients are mostly state owned (tenders)



Conclusions

- AI in digital pathology is a promising direction of research and development
- Digitalization takes over the pathology (slowly)
- AI in digital pathology is hard to
  - Develop
  - Test
  - Deploy

We are working on it!