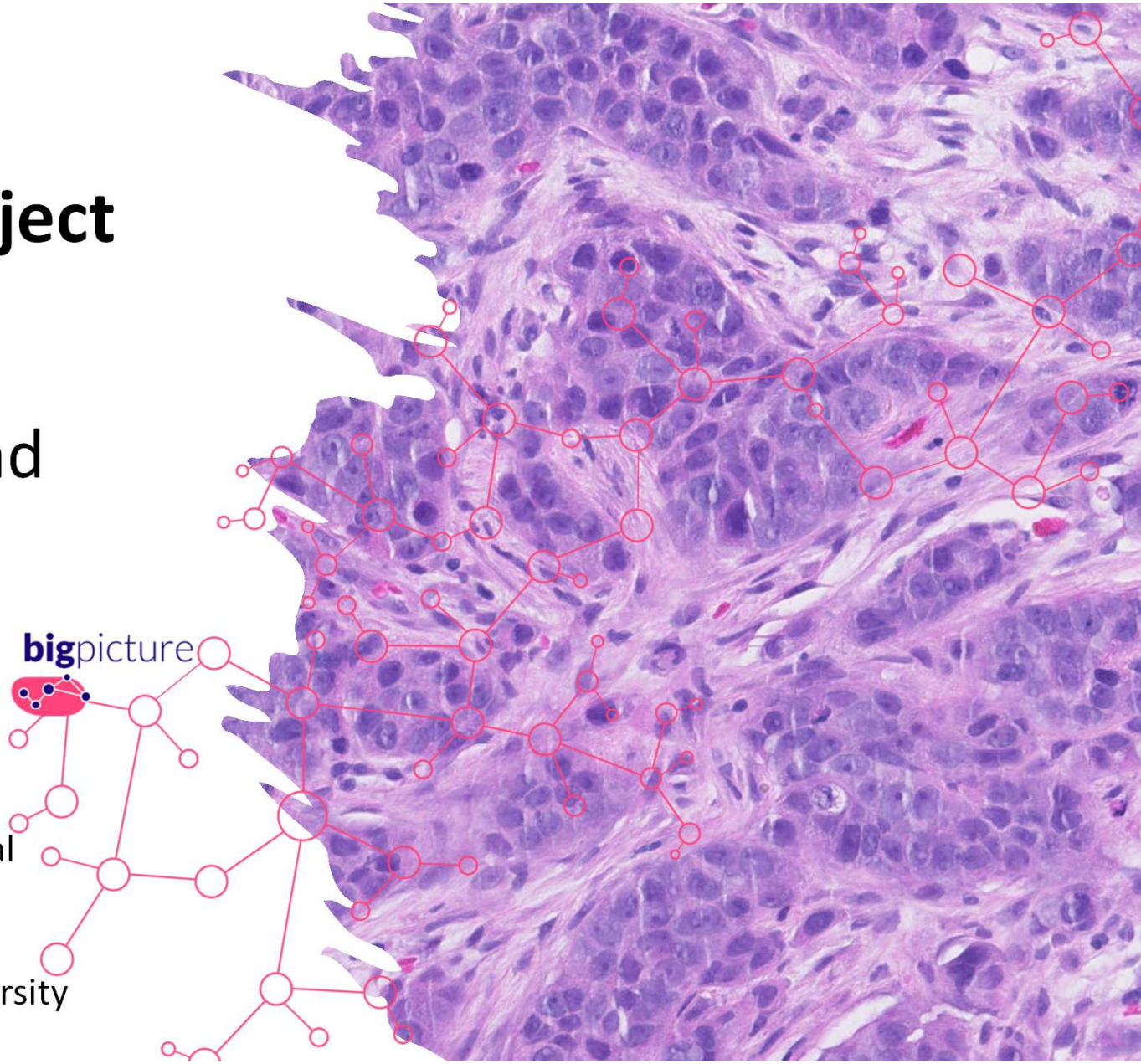
IMI-Bigpicture Project

A Digital Pathology Platform for FAIR Data sharing and AI development

Anna Bodén, WP3 lead and Node Coordinator for Bigpicture

Department of Clinical Pathology
Department of Biomedical and Clinical Sciences
Center of Medical Visualization
Region Östergötland, Linköping University



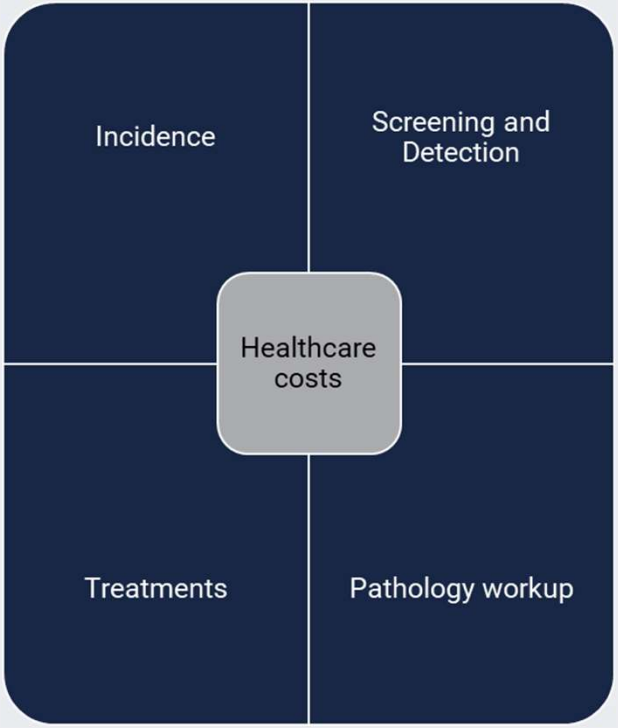
Innovation
Implementation
Clinical Adoption
Research

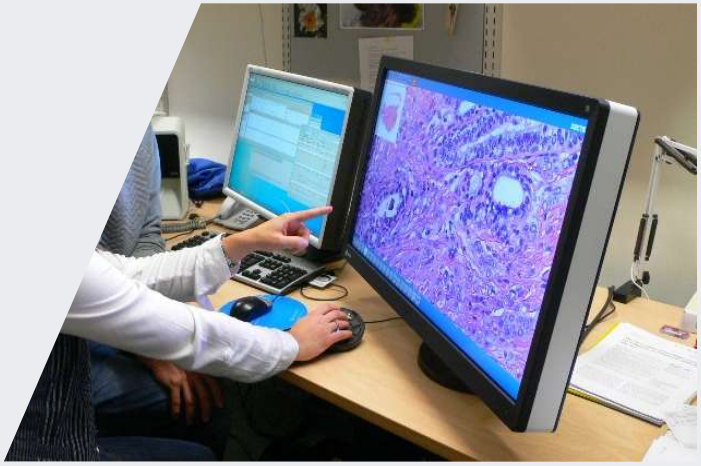
Scanning and Diagnostics	More than 10 years Digital **primary review** > 2TB)
Efficiency and improved patient safety	Macroimages Assembled **cases** Quality **control** **Immediate** access
Improved imaging	Overview Orientation, Alignment Measurements Annotations
Cooperation and collaboration	Consultation Remote access
New workflows	Slides **archived post** scanning Improved **QA**
Improved precision	Image analysis/AI Decision support

LINKÖPING UNIVERSITY

Region Östergötland

Constant changes related to demand and resources





Quantification Ki67 breast HITL

Histopathology

ORIGINAL ARTICLE | Open Access

The human-in-the-loop: An evaluation of pathologists' interaction with AI in clinical practice

Anna C.S Bodén, Jesper Molin, Stina Garvin, Rebecca A. West, Claes Lundström, Darren Treanor

First published: 15 February 2021 | https://doi-org.e.bibl.liu.se/10.1111/his.14356

LiU Full Text

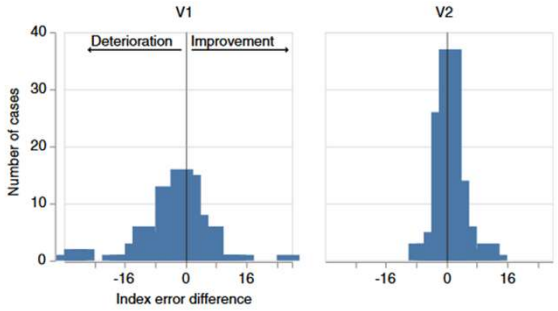
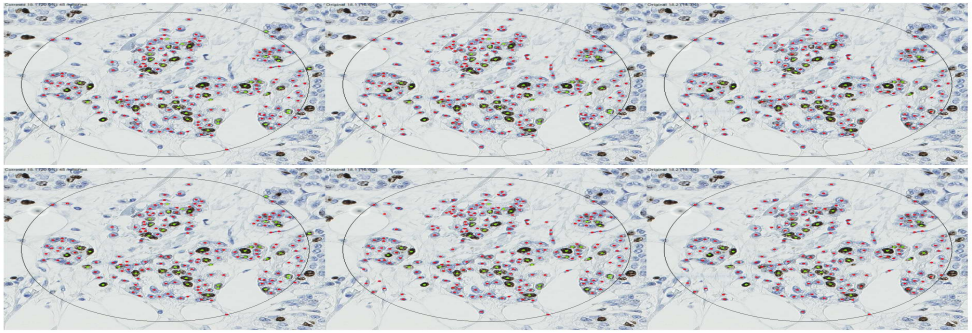


Figure 4. A histogram of case distribution illustrating the difference in Ki67 index error before and after human-in-the-loop correction. A negative value corresponds to a case in which corrections shifted the Ki67 index further away from the ground truth index, and a positive value corresponds to a case in which corrections shifted the Ki67 index towards the ground truth index.



HITL based on V1
4(20.5%, 48 cells modified)

Algorithm V1
(16.5%)

Algorithm V2
(14.3%)

Metastasis Detection Domain Shift

cancers

MDPI

Article

Generalization of Deep Learning in Digital Pathology: Experience in Breast Cancer Metastasis Detection

Sofia Jarkman [1,2,*], Micael Karlberg [2,3], Milda Pocevičiūtė [2], Anna Bodén [1,2], Péter Bándi [3], Geert Litjens [3], Claes Lundström [2,4], Darren Treanor [1,2,5,6] and Jeroen van der Laak [2,3]

[1] Department of Clinical Pathology, and Department of Biomedical and Clinical Sciences, Linköping University, 581 83 Linköping, Sweden
[2] Center for Medical Image Science and Visualization (CMIV), Linköping University, 581 85 Linköping, Sweden
[3] Department of Pathology, Radboud University Medical Center, P.O. Box 9101, 6500 HB Nijmegen, The Netherlands
[4] Sectra AB, Teknikringen 20, 583 30 Linköping, Sweden
[5] Leeds Teaching Hospitals NHS Trust, St James s University Hospital, Beckett Street, Leeds LS9 7TF, UK
[6] Department of Pathology, University of Leeds, Woodhouse Lane, Leeds LS2 9JT, UK
* Correspondence: sofia.jarkman@liu.se

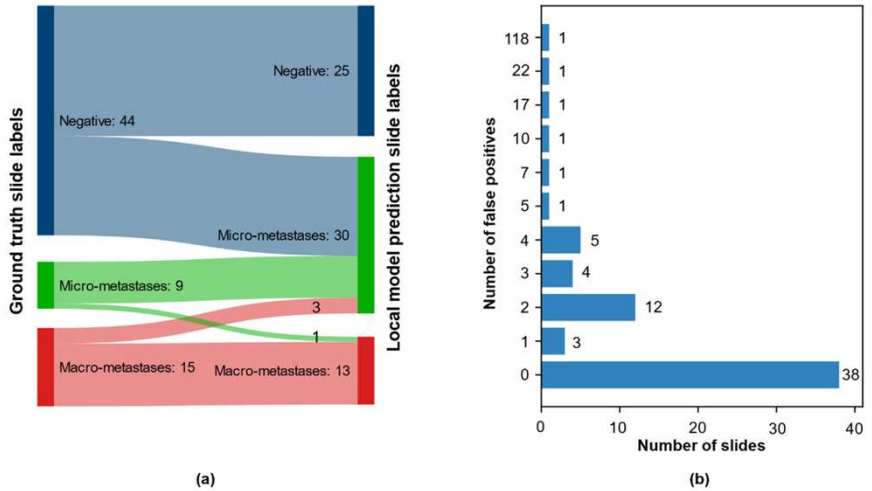
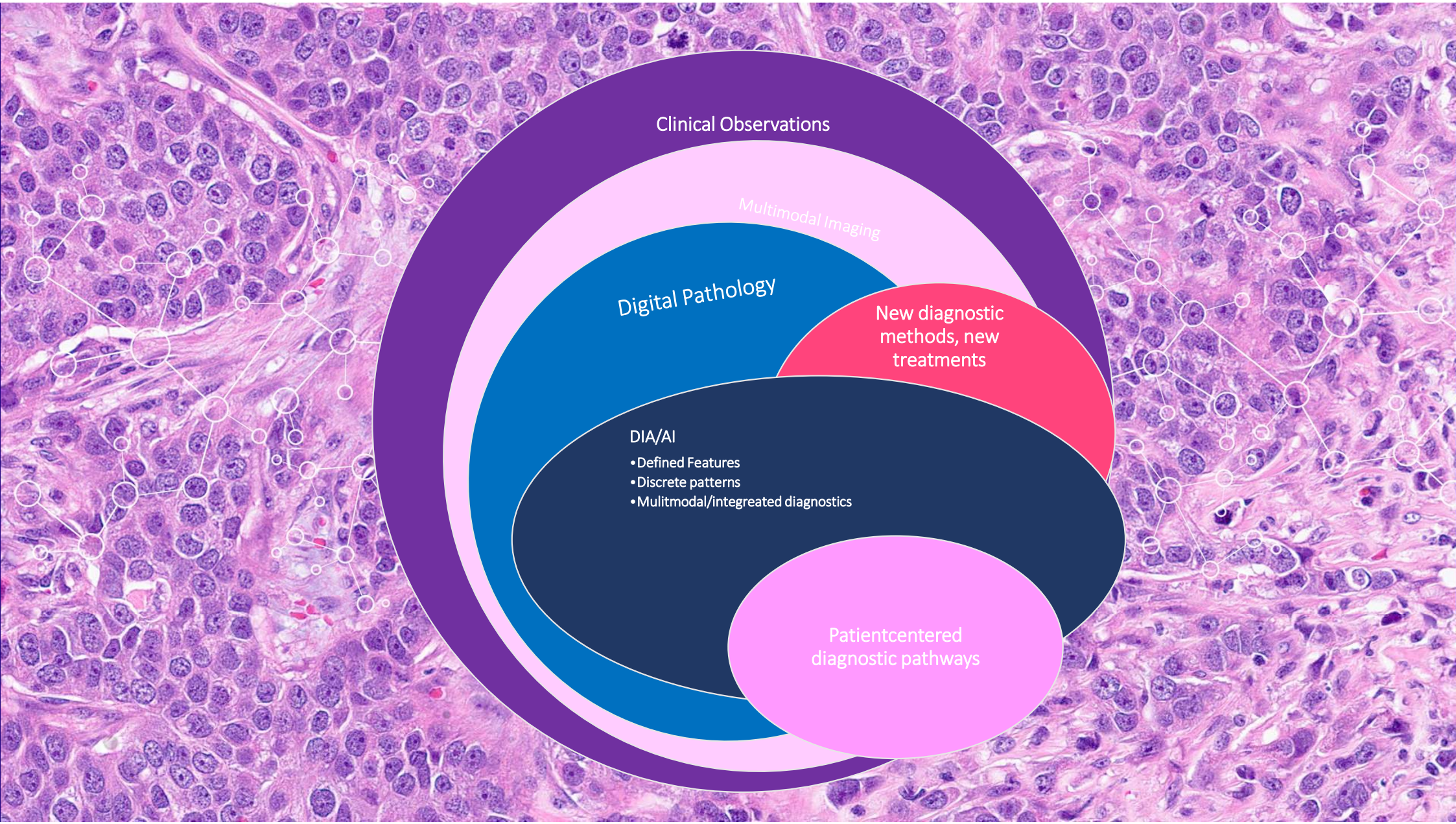


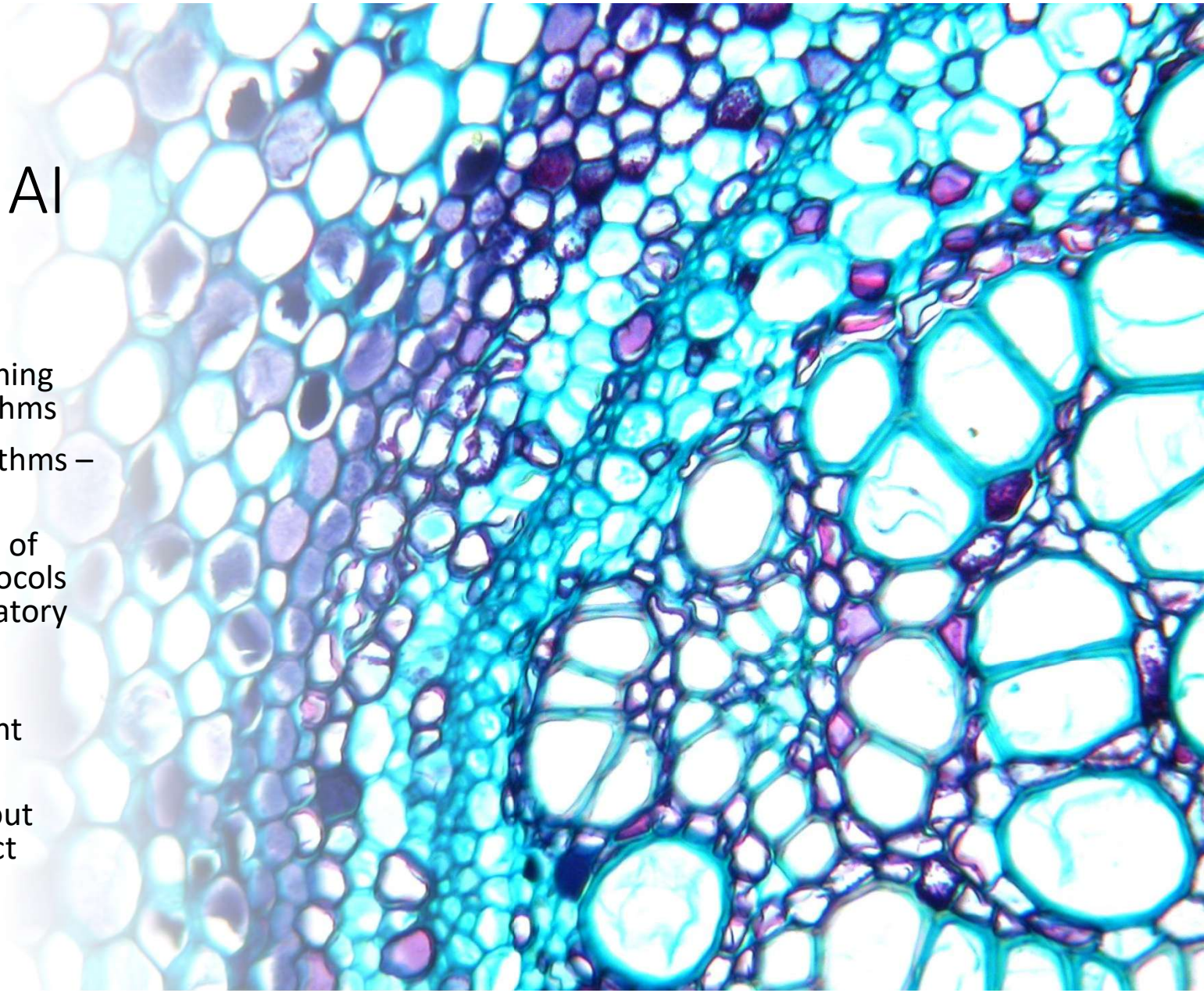
Figure 3. Results from pathologist qualitative evaluation of local model predictions: (a) Sankey diagram over the combined test set (LocalSentinel + LocalAxillary), n = 68, with the ground-truth slide label on left side and local model prediction slide label on right side. Numbers represent the number of slides in each diagnosis group. (One negative slide in ground truth misclassified as negative); and (b) the distribution of false positive across the 68 slides: 38 slides had no false positive; most of the slides with false positives contained 1-4 false positive; a small number of slides had a large number of false positives.

Region Östergötland


Clinical Observations
Multimodal Imaging
Digital Pathology
New diagnostic methods, new treatments
DIA/AI
•Defined Features
•Discrete patterns
•Mulitmodal/integreated diagnostics
Patientcentered diagnostic pathways

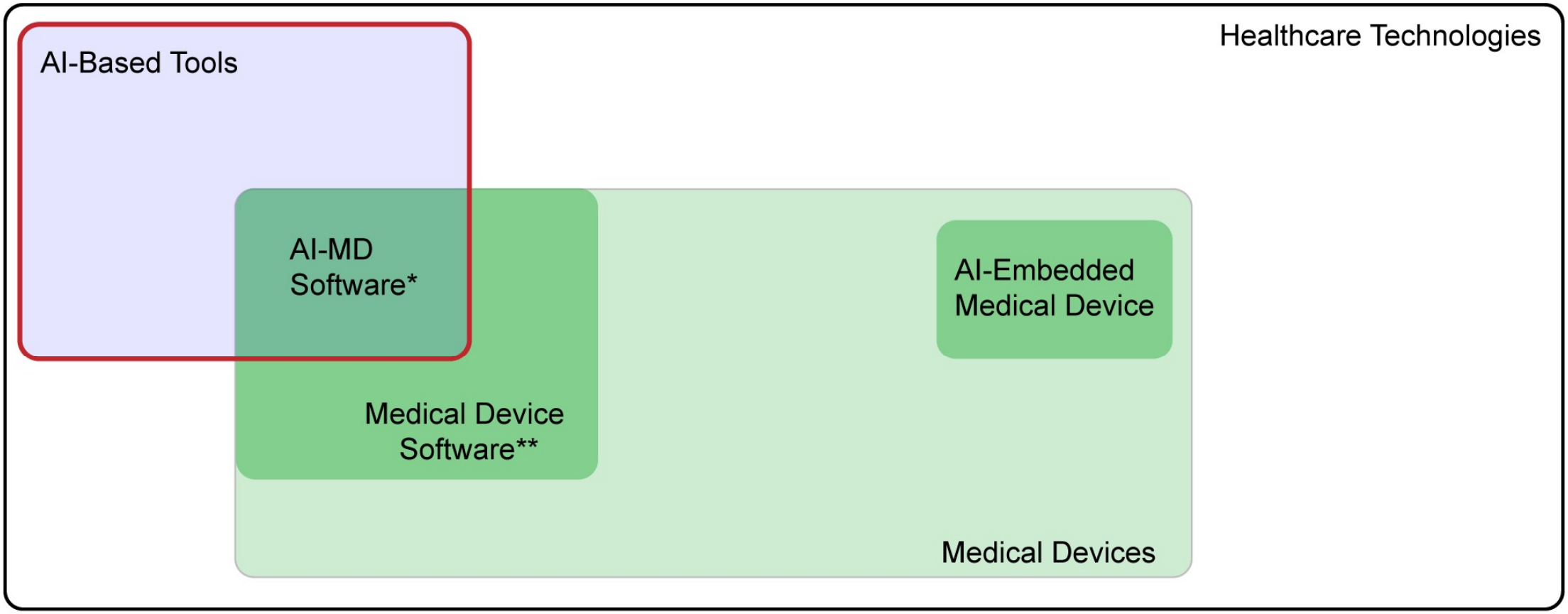
Challenges in implementing AI in clinic

- Deficiencies in data for training and testing/validate algorithms
- Lack of robustness in algorithms – Domain Shift
- Incomplete standardization of image technology and protocols (besides variations in laboratory procedures)
- Demand for improving AI performance to complement pathologists' diagnosis
- Prove patient benefit without displacement/adverse effect



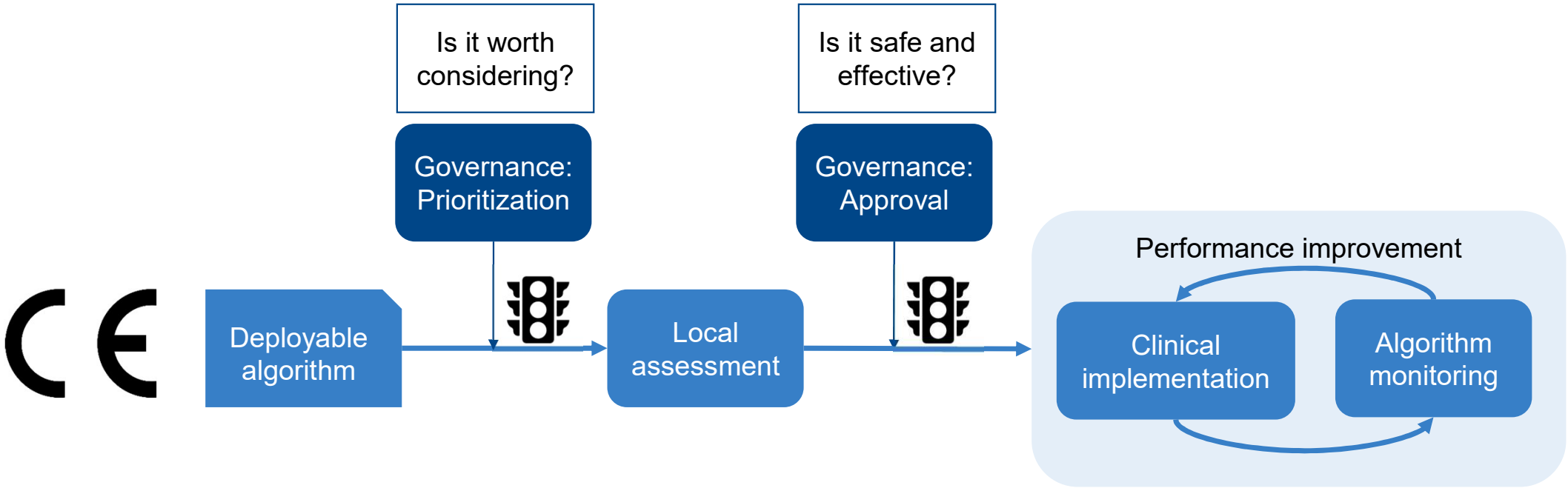
Compliance with legal frameworks

- EU AI Act
  - The EU AI Act categorizes clinical AI as high-risk, meaning hospitals must ensure transparency, explainability, and robust validation before deploying AI tools. This can slow down adoption but ensures safer implementation.
- General Data Protection Regulation (GDPR)
  - Enforces strict data privacy rules, requiring hospitals to obtain patient consent for AI-driven data processing. It also mandates secure data storage and handling, limiting unrestricted AI use.
- EHDS
  - will have a significant impact on digital pathology by improving data accessibility, interoperability, and research opportunities across Europe
- European Medicines Agency (EMA) Guidelines/ Svenska Läkemedelsverket
  - The EMA has developed frameworks for AI use in medicine, including guidance on AI-driven drug development, clinical trials, and regulatory decision-making
  - Läkemedelsverket has a published guidline on AI implementation in Swedish healthcare. It highlights the importance of a systematic, risk-based, and well-documented approach when integrating AI into clinical settings. The guideline also includes a checklist to support healthcare providers in planning AI adoption.
    - You can explore the full guideline here.
- Medical Device Regulation (MDR)
  - AI-powered clinical tools often fall under MDR, requiring rigorous testing, validation, and certification before deployment.
- Ethical AI Principles (EU Ethics Guidelines for Trustworthy AI, WHO Guidance on Ethics and Governance of Artificial Intelligence for Health)
  - Hospitals must ensure AI decisions align with ethical guidelines, preventing bias, ensuring fairness, and maintaining patient trust. Fostering responsibility and explainability as well as protect human autonomy.



* AI/ML-enabled medical device software (US) and AI medical device software (EU)
** Software as a medical device (US) and Medical device software (EU)

Region Östergötland



Region Östergötland

Challenges for AI development in digital pathology

- The lack of available data for AI development in digital pathology makes it difficult to develop new algorithms.
- The lack of data can lead to poorer AI performance.
- Collaboration and data sharing can help overcome the lack of data for AI development in digital pathology (and comply to EHDS)
- EHDS
  - will have a significant impact on digital pathology by improving data accessibility, interoperability, and research opportunities across Europe

Bigpicture Values

Trustworthy

Covering every detail by legal & ethical frameworks, ensuring privacy and reliability.

Catalyzing

Accelerating AI in pathology to benefit patients by creating innovative solutions for research.

Collaborative

Bridging gaps and fostering innovation through partnerships.

Inclusive

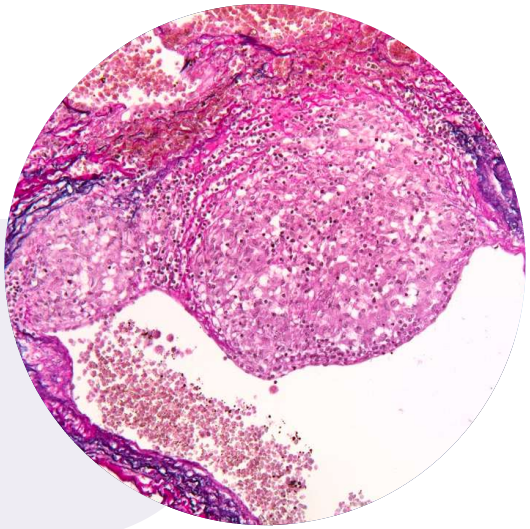
Building a diverse community including pathologists, pharmaceutical companies, AI researchers, policy makers, clinicians & patients.

bigpicture

Bigpicture facts & figures

- 44 partners
- 15 countries
- € 32.319.825 IMI funding
- € 37.762.082 EFPIA funding
- Start: 1 February 2021
- Duration: 6 years

2 yrs to go

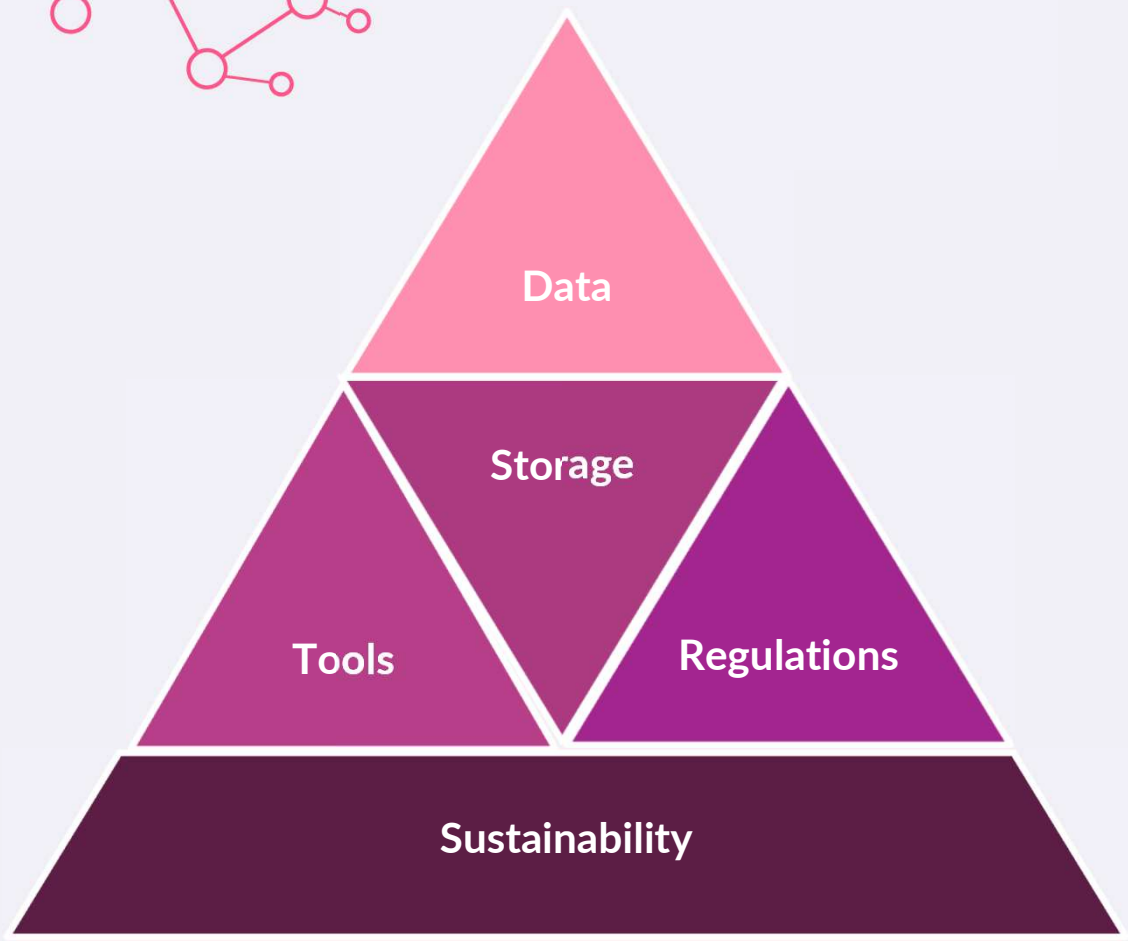


bigpicture

Sustainability by design

- ✓ **Community-based approach**
- ✓ **Bi-directional value stream**
  - ✓ for contributors and users
  - ✓ for non clinical and clinical partners
    - ✓ *advantages of transfer learning from domain specific model*
- ✓ **Platform where functionality and data co-exist**
- ✓ **Platform will be dynamic; growth planned**

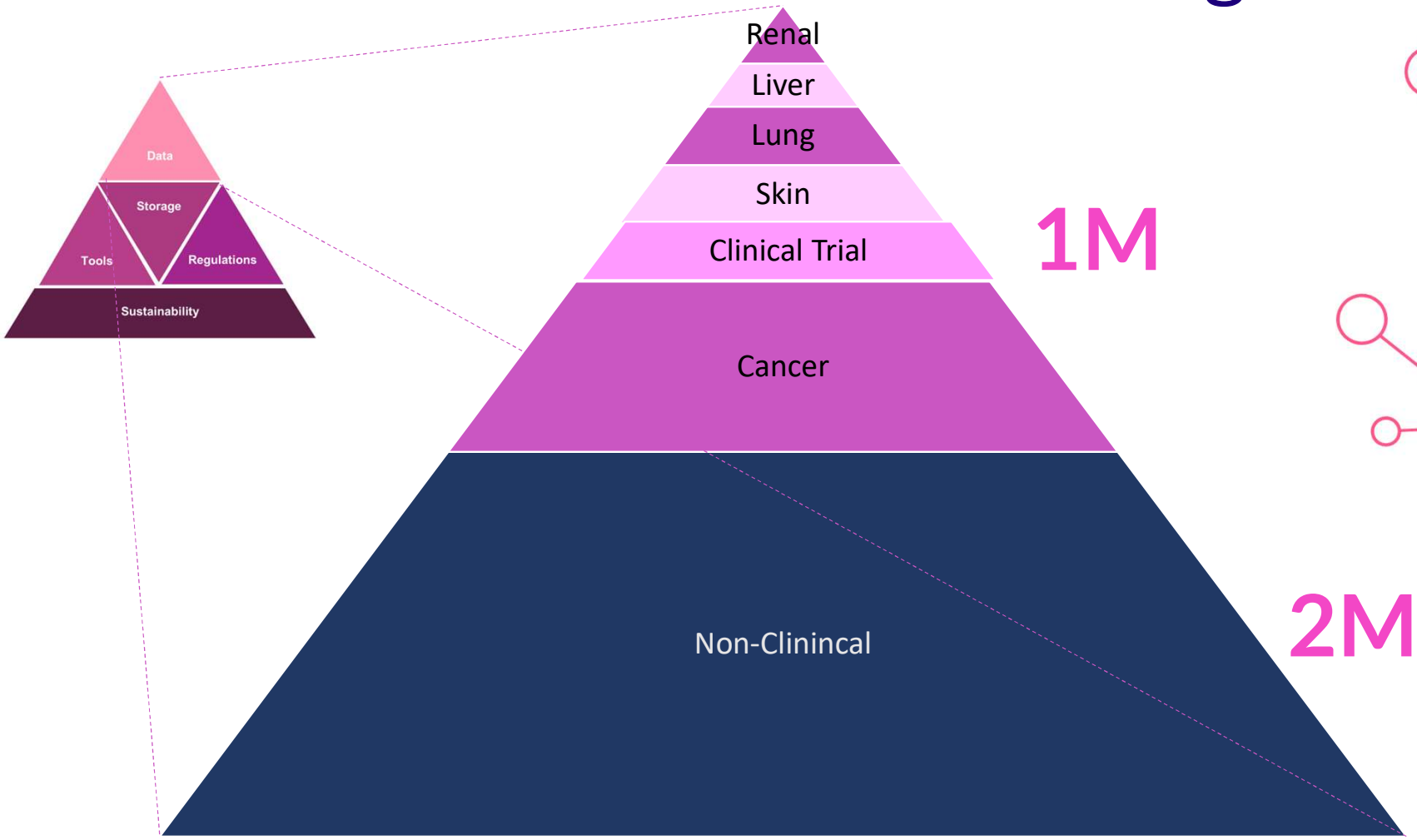
bigpicture



Project structure

WP3	Data: 3M digital slides
WP2	Storage/infrastructure: 4.5 Pbytes
WP4	Generic tools: submission, access, analysis and AI models, viewer
WP5	Regulations: acceptance, usage sharing
WP6	Sustainable platform

3 million whole slide images



bigpicture

Lead and Core objectives WP3

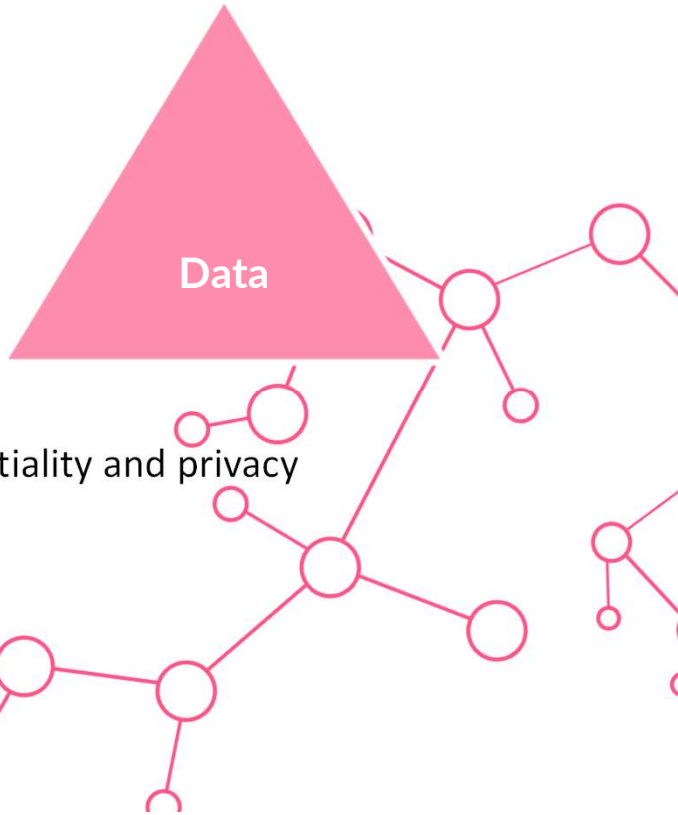

Renate Kain


Anna Boden


Brian Knight
Boehringer Ingelheim


Erio Barale-Thomas
Janssen R&D

1. **A sustainable node network**
   - Mediating and collecting datasets from beneficiaries and third-parties
   - Contribute them to the BIGPICTURE data repository
2. **Datasets** prioritised within each node from different sources
   - Support WP4 for developing innovative tools and algorithms
3. **Data collection of highest quality**
   - Prerequisites: conversion, standardisation and harmonisation
   - Using the WP2 infrastructure
4. **Security mechanisms** for optimal management
   - Honest broker: fair, regulated data access and sharing, maintaining confidentiality and privacy



bigpicture

Node coordination clinical network

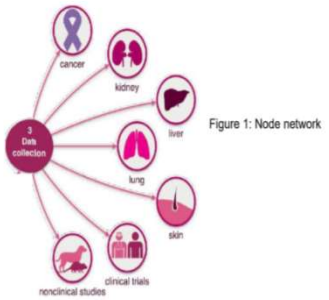
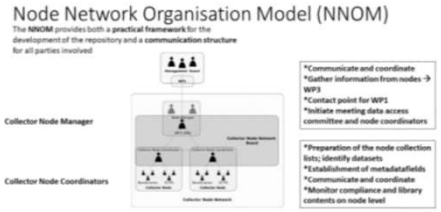
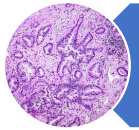
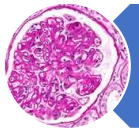


Figure 1: Node network

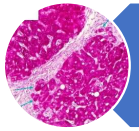




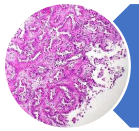
Cancer node: Paul van Diest, Utrecht



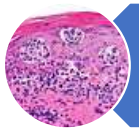
Kidney/Renal node: Renate Kain, Vienna



Liver node: Darren Treanor, Leeds



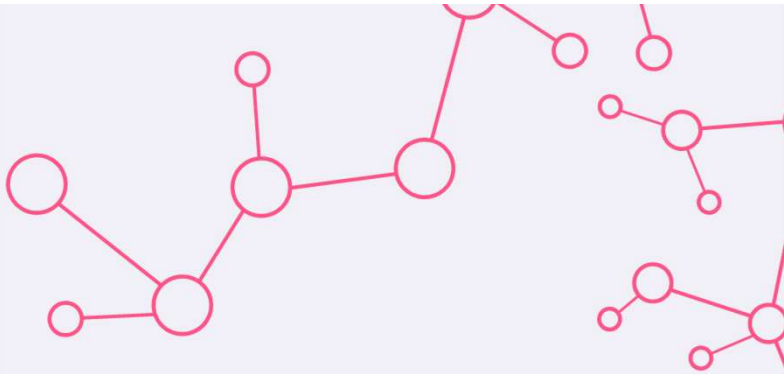
Lung node: Ollie Carpen, Helsinki



Skin node: Anna Bodén, Linköping

Clinical trial node: Carsten Denkert, Marburg

Bigpicture Platform
4 use case scenarios



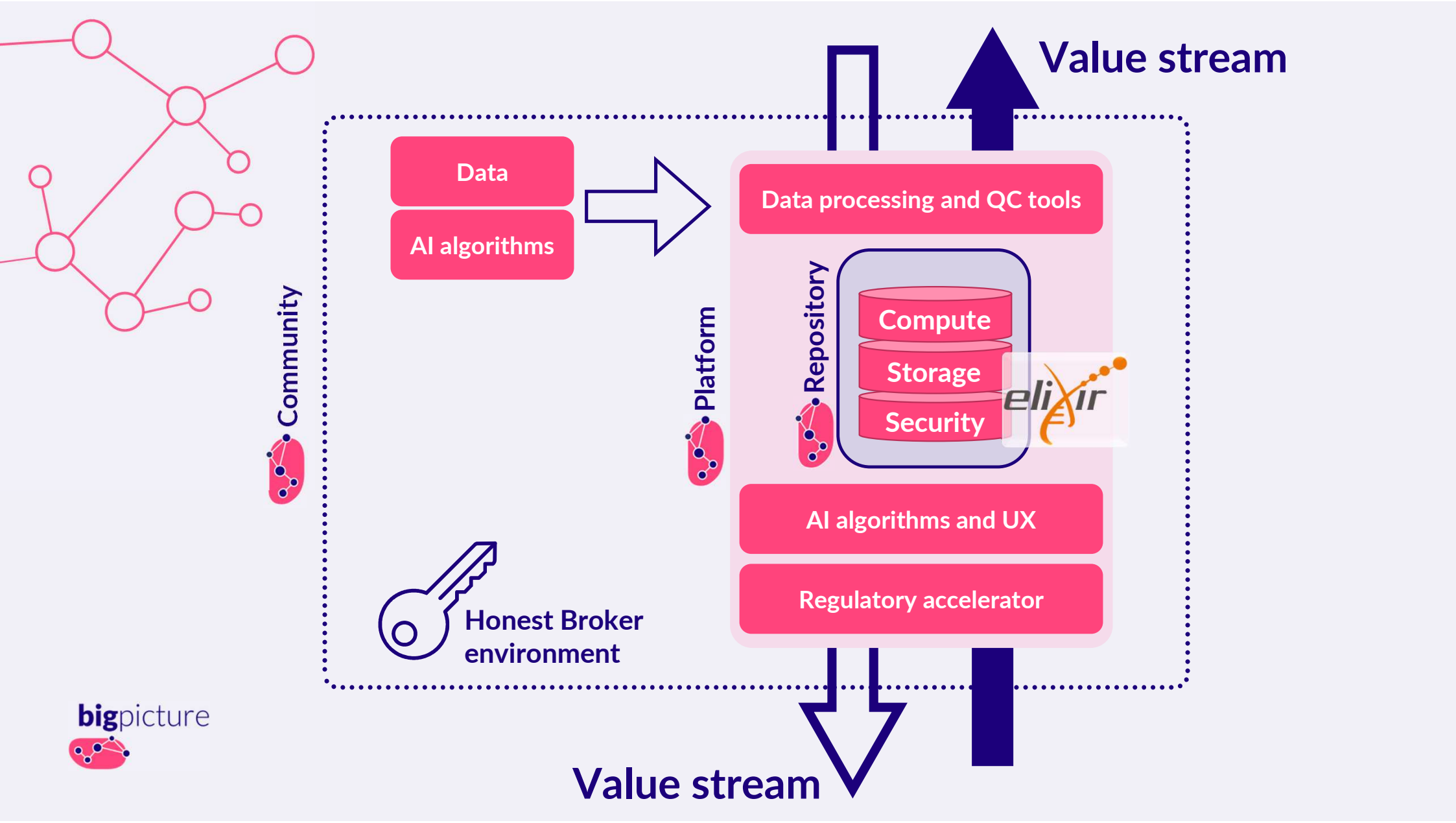
Data upload
Share data
Run algorithm and receive result

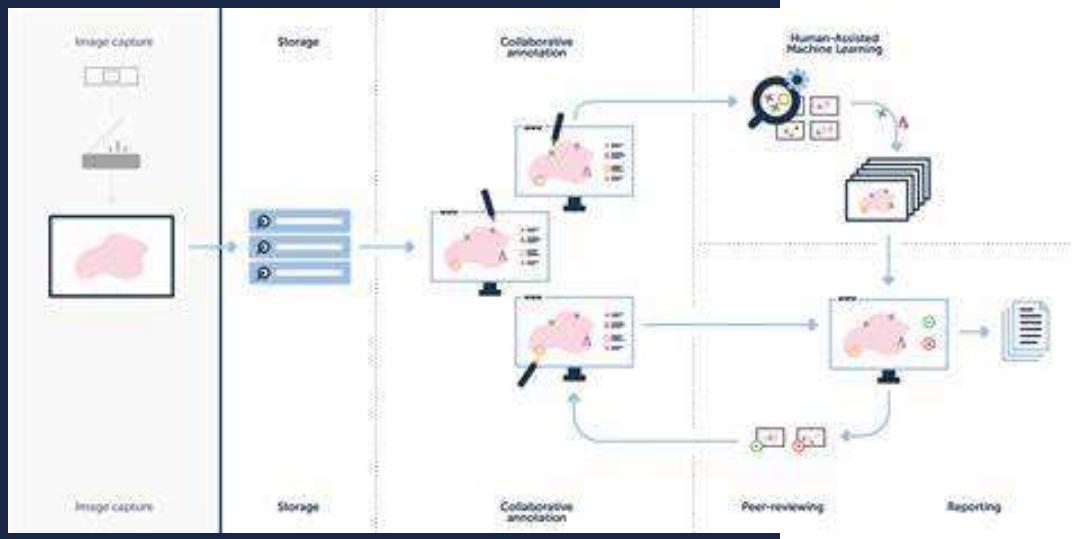
Data download
to develop algorithm

Algorithm download
Run algorithm and receive result

Algorithm upload
to receive results

bigpicture


Value stream
Data
AI algorithms
Data processing and QC tools
Community
Platform
Repository
Compute
Storage
Security
elixir
AI algorithms and UX
Regulatory accelerator
Honest Broker environment
bigpicture
Value stream

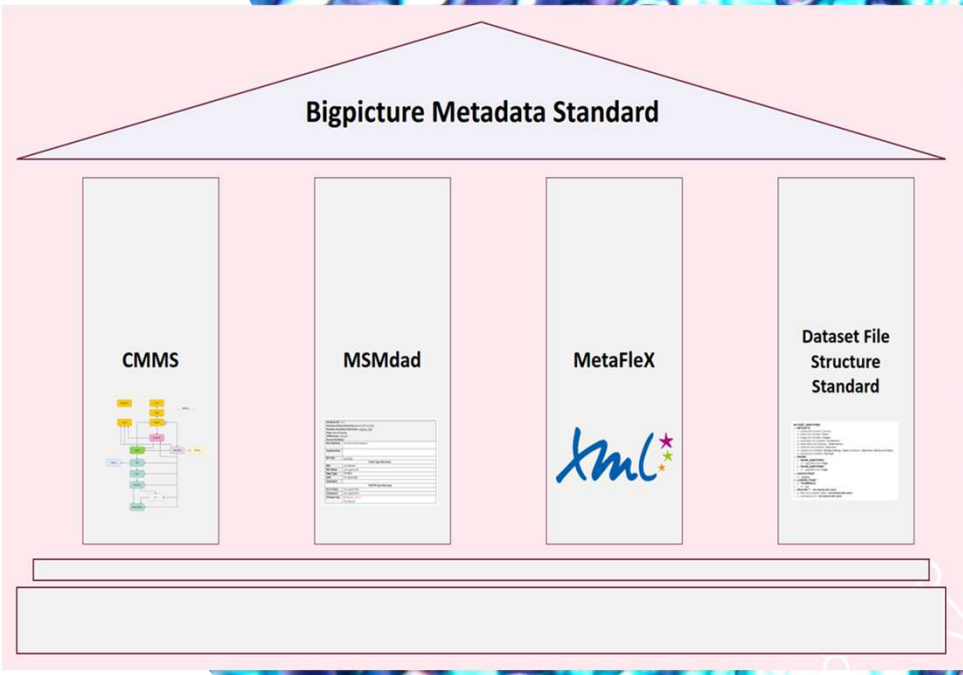


- Integration of Cytomine **web viewer** (open source) will be the graphical user interface that can be used on the Bigpicture Platform.
- Cytomine Bigpicture Edition is being extended to serve as the **main user interface** for viewing and annotating images from the BigPicture repository including tailored modules specifically developed for this project.
- It aims to offers features like **multi-image comparison, annotation tools, and real-time collaboration**

— GitHub - Cytomine-ULiege/bigpicture-cytomine-web-ui: Cytomine-Web-UI is the graphical user interface for Cytomine bigpicture edition.

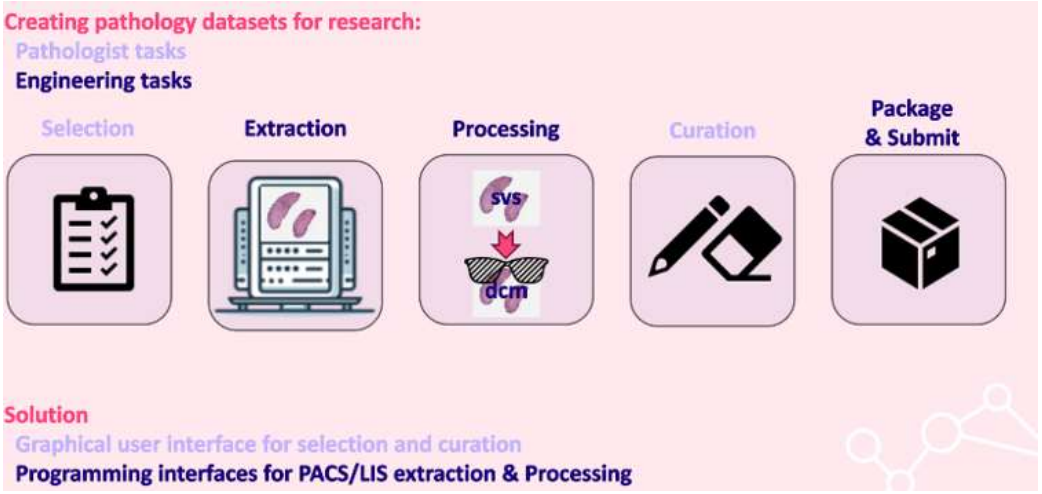
Data and standards

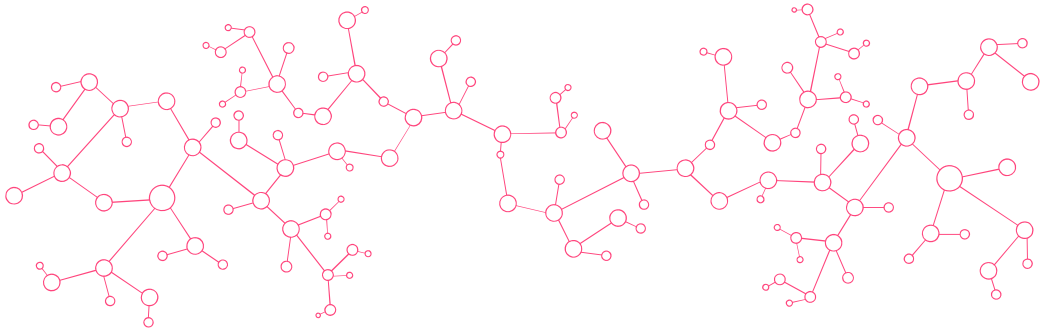
- Metadata and standards enable efficient management and sharing of digital pathology data
- Standards such as DICOM enable interoperability
- Bigpicture use DICOM for WSI and a (soon to be) public metadata model (xml) to enable sharing
- Data from the primary source are converted and curated and quality controlled before submission
- Bigpicture Dataset are either anonymized or pseudonymized, in the latter prospective addition of metadata to a dataset is possible.
- For clinical datasets an ethical approval should be provided
- Sharing data will be GDPR compliant and assisted by a community DSA



bigpicture

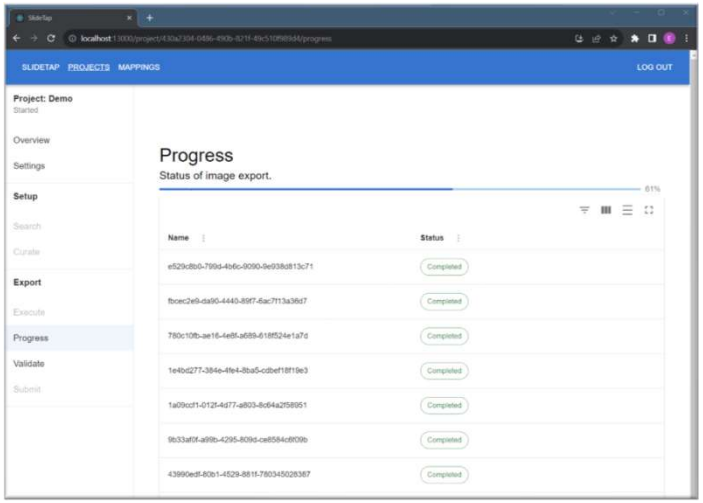
Dataset collection

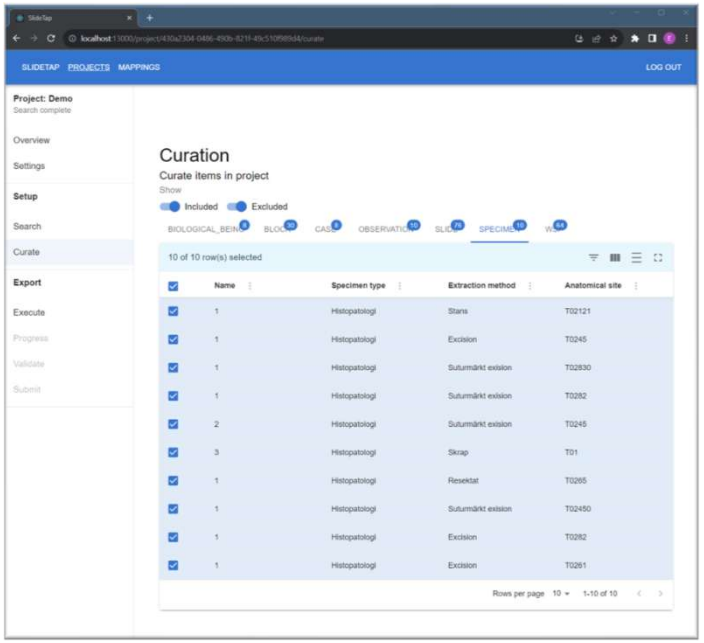


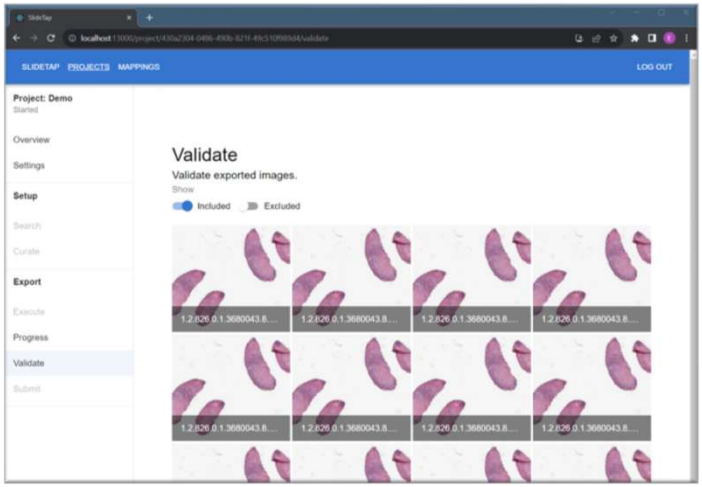


bigpicture

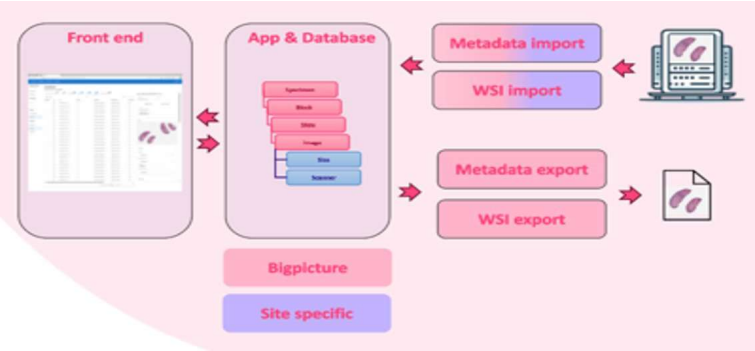
https://github.com/imi-bigpicture/slidetap

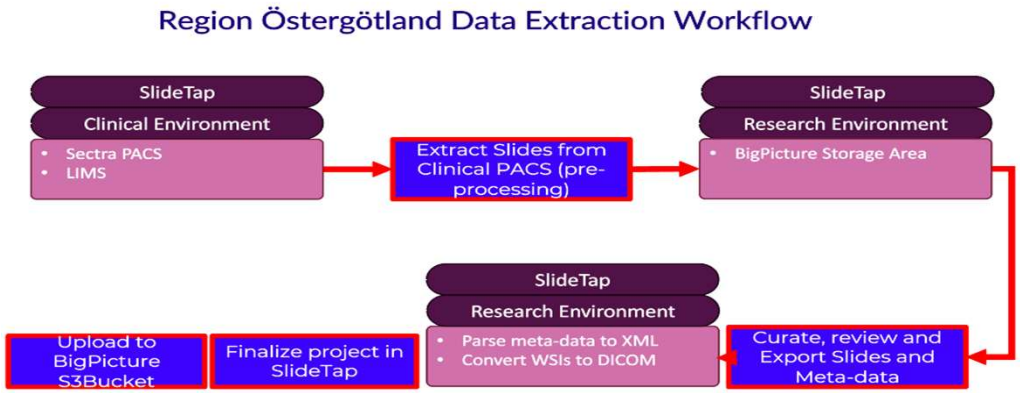


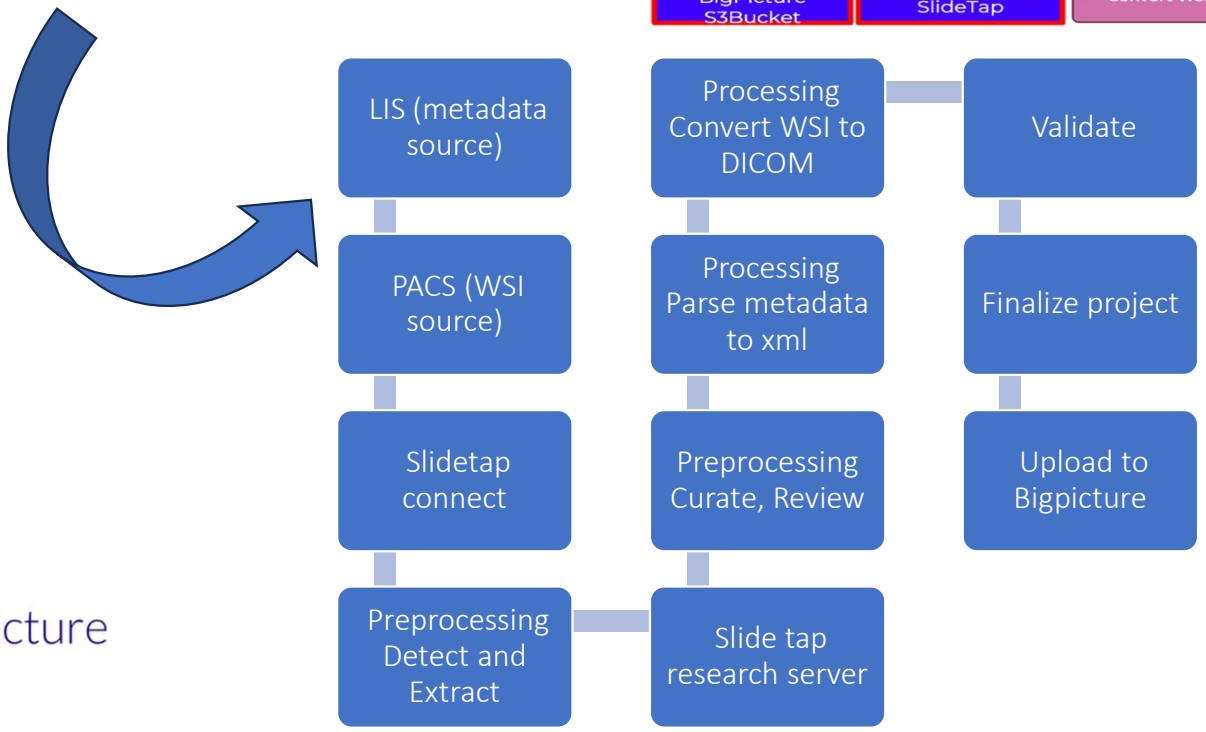




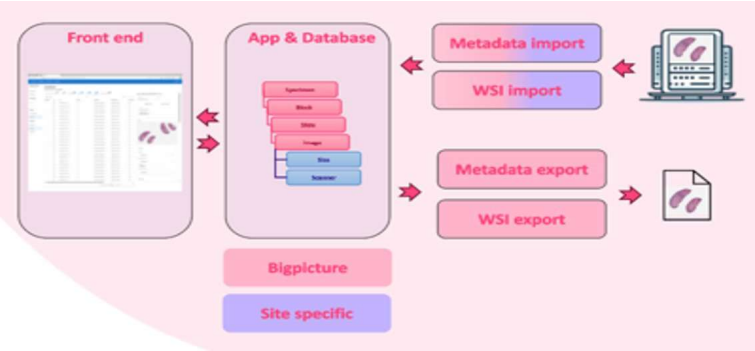
bigpicture

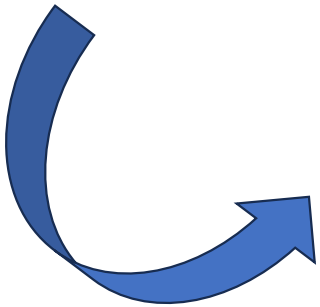

Front end
App & Database
Metadata import
WSI import
Metadata export
WSI export
Bigpicture
Site specific

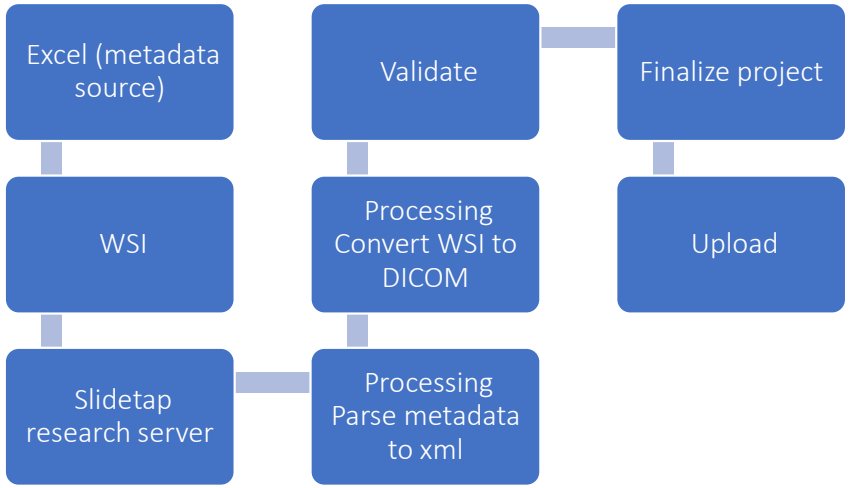

Region Östergötland Data Extraction Workflow
SlideTap
Clinical Environment
Sectra PACS
LIMS
Extract Slides from Clinical PACS (pre-processing)
SlideTap
Research Environment
BigPicture Storage Area
Curate, review and Export Slides and Meta-data
SlideTap
Research Environment
Parse meta-data to XML
Convert WSIs to DICOM
Finalize project in SlideTap
Upload to BigPicture S3Bucket


LIS (metadata source)
PACS (WSI source)
Slidetap connect
Preprocessing Detect and Extract
Slide tap research server
Preprocessing Curate, Review
Processing Parse metadata to xml
Processing Convert WSI to DICOM
Validate
Finalize project
Upload to Bigpicture

bigpicture

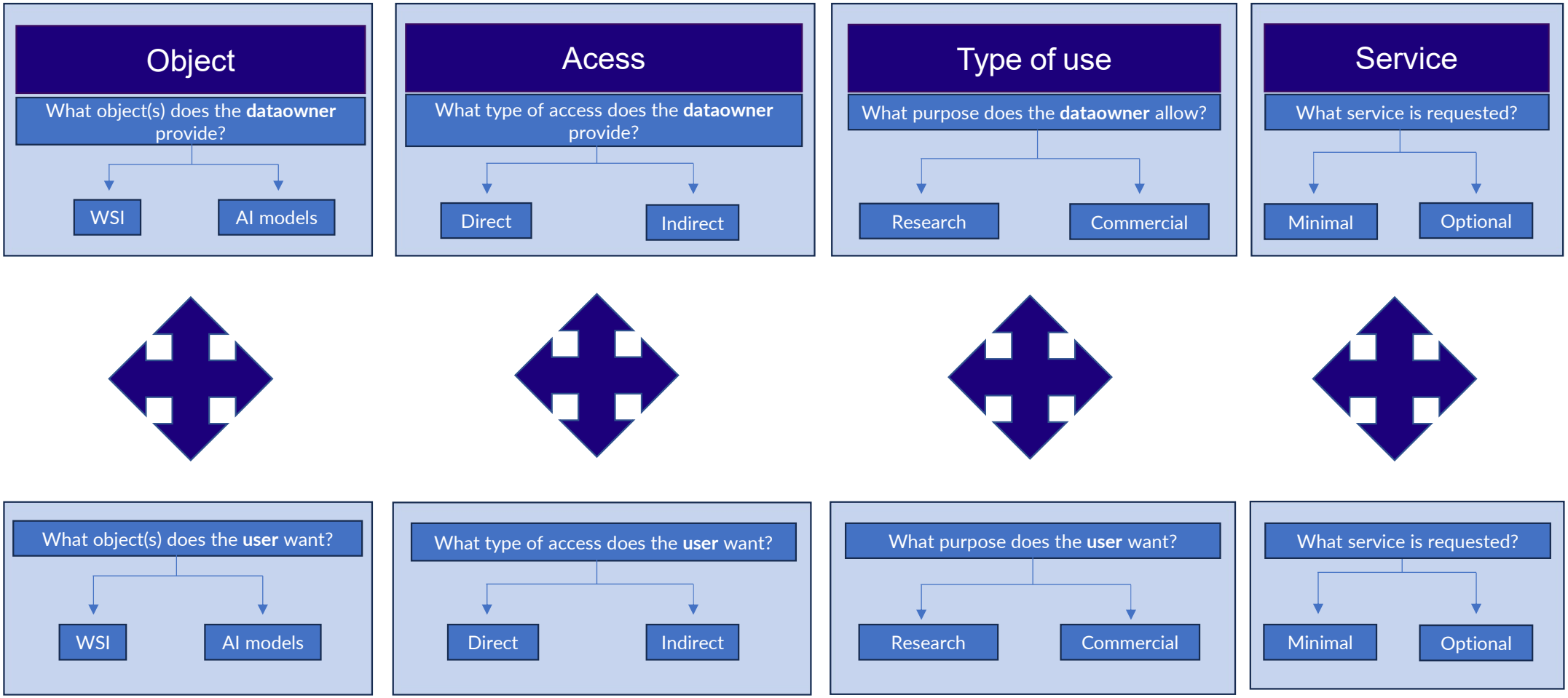

Front end
App & Database
Metadata import
WSI import
Metadata export
WSI export
Bigpicture
Site specific




Excel (metadata source)
Validate
Finalize project
WSI
Processing Convert WSI to DICOM
Upload
Slidetap research server
Processing Parse metadata to xml

bigpicture

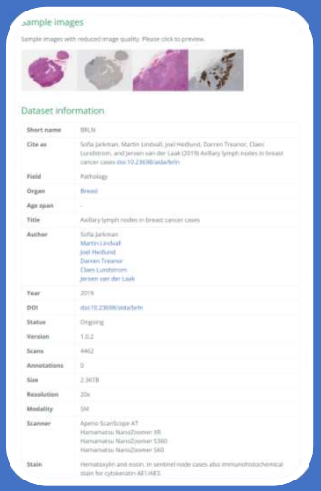
Four dimensions to define ToU for the contributor/user

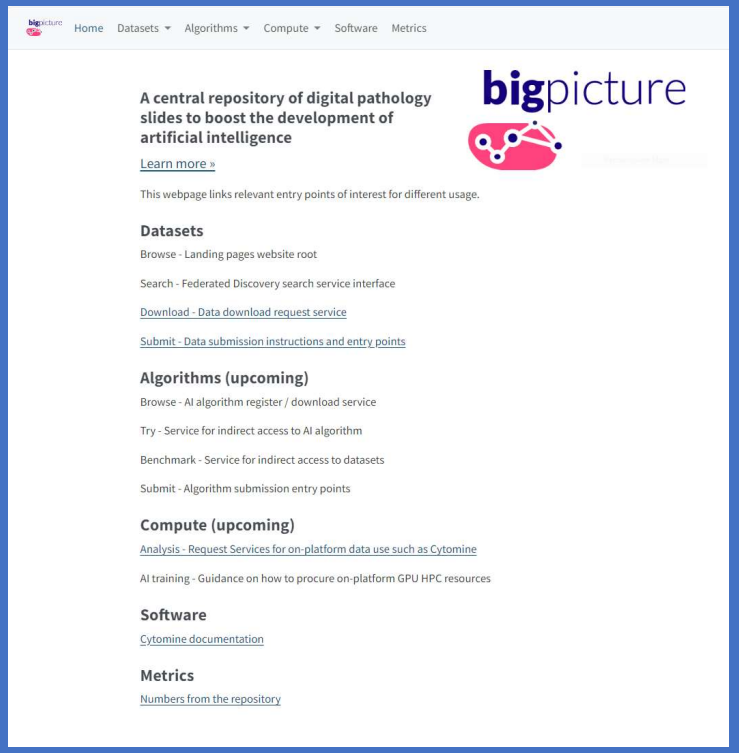


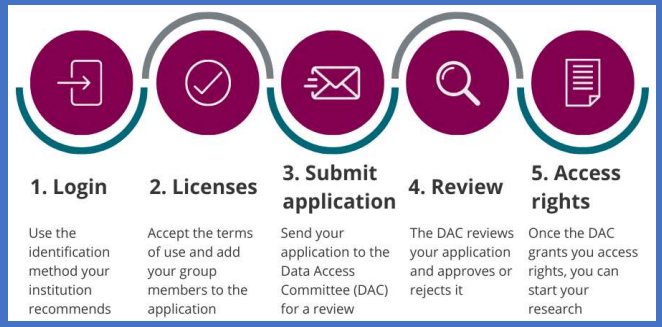
Findable

bigpicture

Browse and search





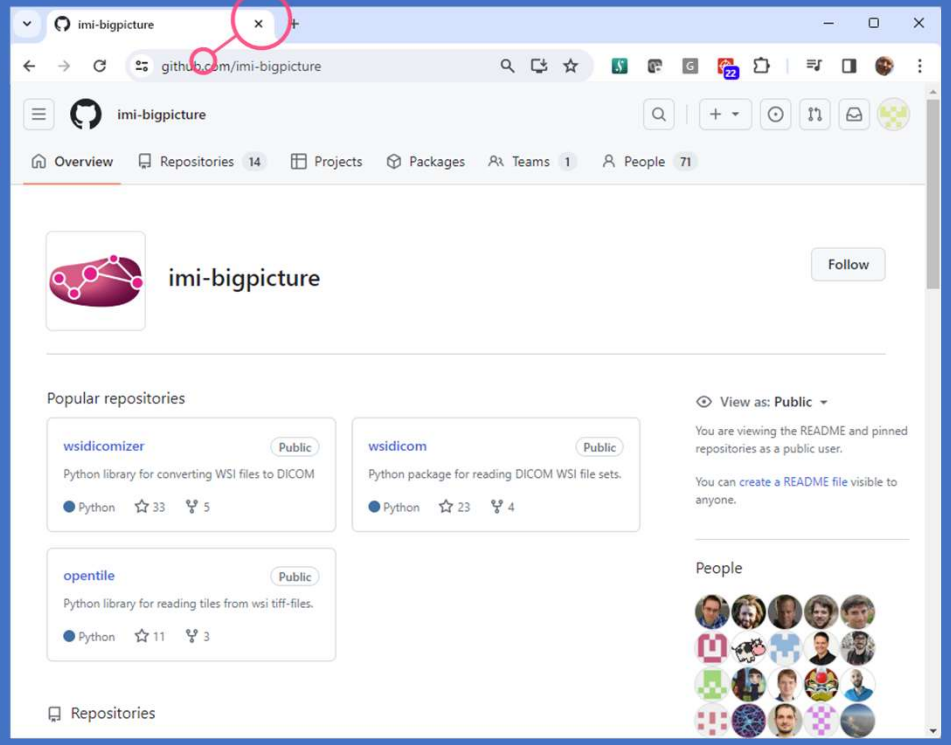


Authentication and Authorization Infrastructure

Resource Entitlement Management System
(AAI – REMS)

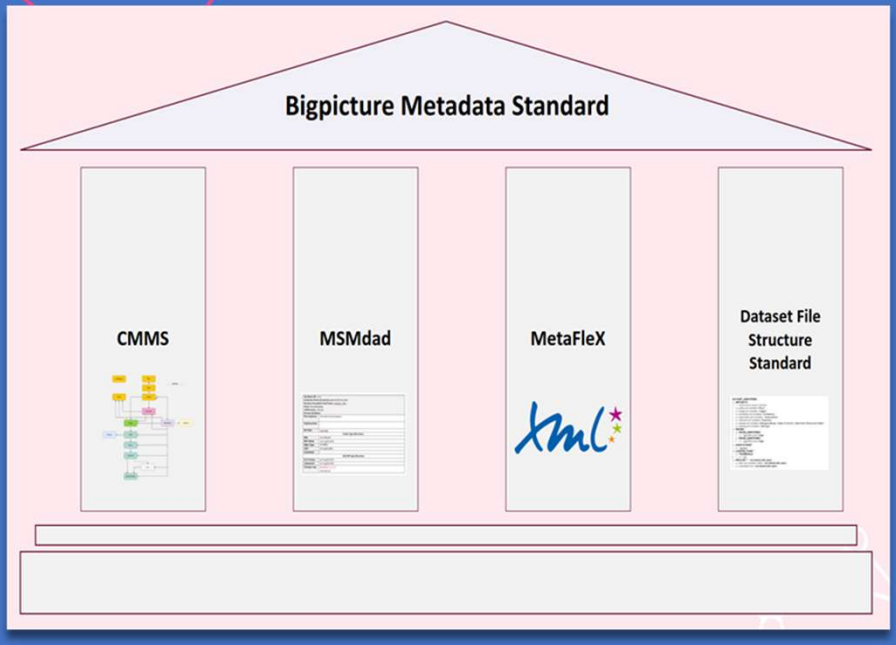
bigpicture

Interoperable



bigpicture

Reusable



bigpicture

Construction of an extensive human skin dataset for artificial intelligence development

Jeronimo Frias Rose[1], Anna Bodén[1,2] Caroline Bivik Stadler[3] and Jeroen van der Laak[2,3].
[1] Department of Clinical Pathology, and Department of Biomedical and Clinical Sciences, Linköping University, Linköping, Sweden, [2] Center for Medical Image Science and Visualization (CMIV), Linköping University, Linköping, Sweden, [3] Department of Pathology, Radboud University Medical Center, 6525 GA Nijmegen, The Netherlands.

Background & objectives

The Bigpicture (BP) consortium is a central repository of digital pathology slides to boost the development of artificial intelligence, it consists of members from both private and public organizations. BP's main goal is to create the first European General Data Protection Regulation (GDPR) compliant platform, where quality-controlled whole slide images (WSI) and advanced AI algorithms will co-exist

Methods

The department of clinical pathology in Region Östergötland has a digital image archive of >2 Petabytes. In order to help achieve BP's goal of 3 million WSI with their associated metadata, we chose to participate as contributors to the repository. We developed a protocol for dataset extraction, that complies with all applicable regulations, ensuring high quality content.

Results

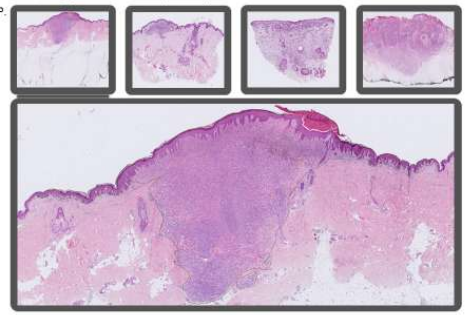
A dataset of approximately 45,000 WSI with their associated metadata was compiled, We were tasked with gathering WSI from skin samples. A human skin dataset was designed, mirroring daily-basis clinical cases and their associated WSI as our contribution to the BP repository. After ethical approval for using patient data for research, we selected skin cases from patients 18 years and older, from 2019-2022, including cases with only one diagnosis amongst melanoma, other melanocytic lesions, squamous cell carcinoma, basal cell carcinoma, dermatofibroma, seborrheic keratosis, actinic keratosis and scar tissue, the diagnoses were organized in different groups (Table 1). Metadata was partly preserved (patient age, anatomical site, acquisition time, laboratory related data and diagnosis/observations). All data anonymization, conversion and extraction was automated, using tools that were specifically developed within BP.

Selection of diagnoses for the dataset

Diagnosis	Group
Malignant melanoma of skin (disorder)	Malignant melanocytic
Melanoma in situ of skin (disorder)	Dysplastic lesions and in situ melanoma
Dysplastic nevus of skin (disorder)	Dysplastic lesions and in situ melanoma
Melanocytic nevus of skin (disorder)	Benign melanocytic lesions
Blue nevus of skin (disorder)	Benign melanocytic lesions
Epithelioid and spindle cell nevus (disorder)	Benign melanocytic lesions
Epithelioid and spindle cell nevus (disorder)	Benign melanocytic lesions
Basal cell carcinoma of skin (disorder)	Basal cell carcinoma
Squamous cell carcinoma of skin (disorder)	Squamous cell carcinoma
Squamous cell carcinoma in situ of skin (disorder)	Squamous cell carcinoma in situ
Seborrheic keratosis (disorder)	Benign skin lesions
Actinic keratosis (disorder)	Benign skin lesions
Scar of skin (disorder)	Benign skin lesions
Dermatofibroma (disorder)	Benign skin lesions

Conclusion

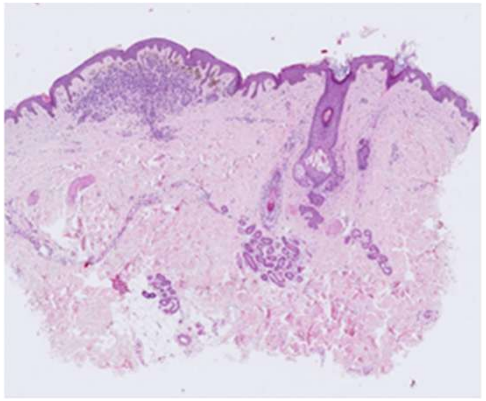
The main limiting factor when developing and implementing AI-tools is availability of data, which can be attributed to challenges with data quality, storage and regulations for patient data protection, and BP helped us overcome said challenges. We succeeded in the compilation of an extensive clinically relevant dataset for BP's repository, which will be useful for research purposes and development of relevant AI solutions. The increasing adoption of digital pathology is an enabler for the development of AI-based tools that support histopathological diagnostics, thus leading to more accurate diagnoses and improvement in patient care.

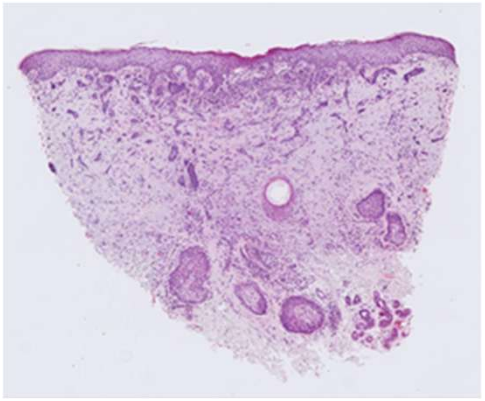


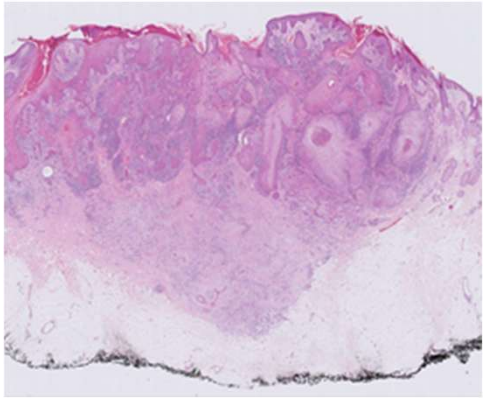
bigpicture.eu/

LINKÖPING UNIVERSITY
CENTER FOR MEDICAL IMAGE SCIENCE
AND VISUALIZATION, CMIV

Jeronimo.Frias.Rose@regionostergotland.se
Region Östergötland

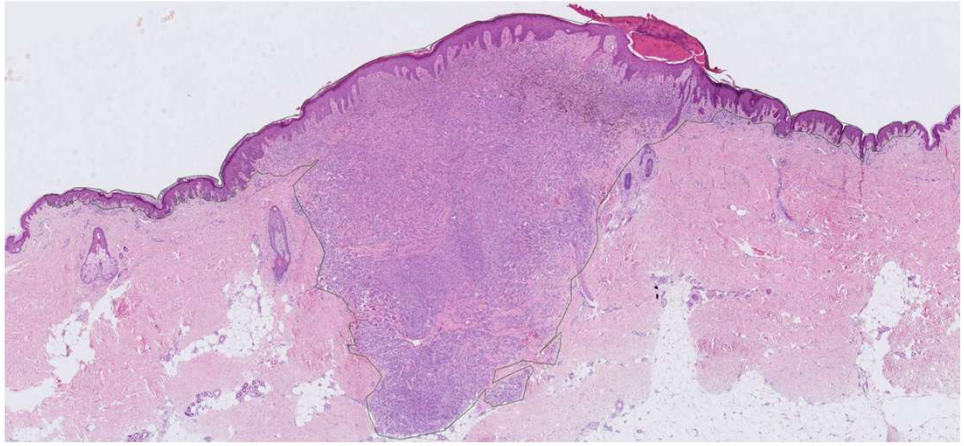






Selection of diagnoses for the dataset

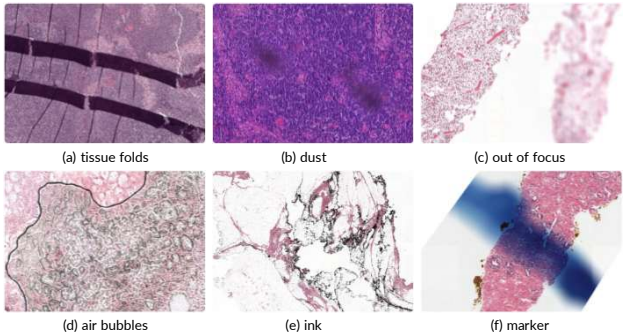
Diagnosis	Group
Malignant melanoma of skin (disorder)	Malignant melanocytic
Melanoma in situ of skin (disorder)	Dysplastic lesions and in situ melanoma
Dysplastic nevus of skin (disorder)	Dysplastic lesions and in situ melanoma
Melanocytic nevus of skin (disorder)	Benign melanocytic lesions
Blue nevus of skin (disorder)	Benign melanocytic lesions
Epithelioid and spindle cell nevus (disorder)	Benign melanocytic lesions
Epithelioid and spindle cell nevus (disorder)	Benign melanocytic lesions
Basal cell carcinoma of skin (disorder)	Basal cell carcinoma
Squamous cell carcinoma of skin (disorder)	Squamous cell carcinoma
Squamous cell carcinoma in situ of skin (disorder)	Squamous cell carcinoma in situ
Seborrheic keratosis (disorder)	Benign skin lesions
Actinic keratosis (disorder)	Benign skin lesions
Scar of skin (disorder)	Benign skin lesions
Dermatofibroma (disorder)	Benign skin lesions



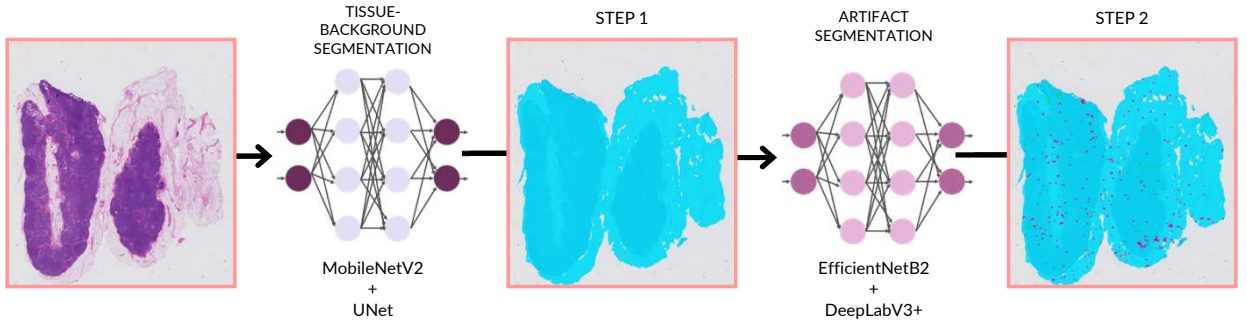
Artifact segmentation and quality control

bigpicture

ARTIFACTS

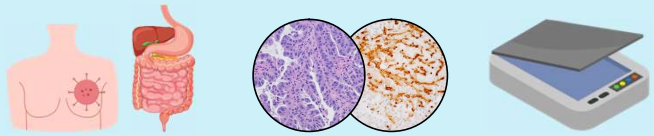


MODEL



DATA

TRAINING DATA: 100 cases



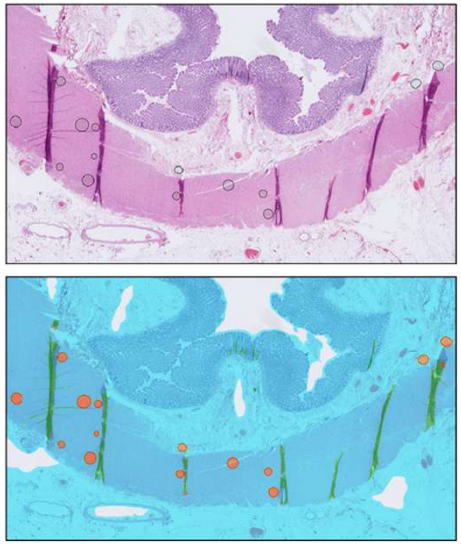
9 tissue types | 9 stains | 7 scanners

VALIDATION: 500 cases

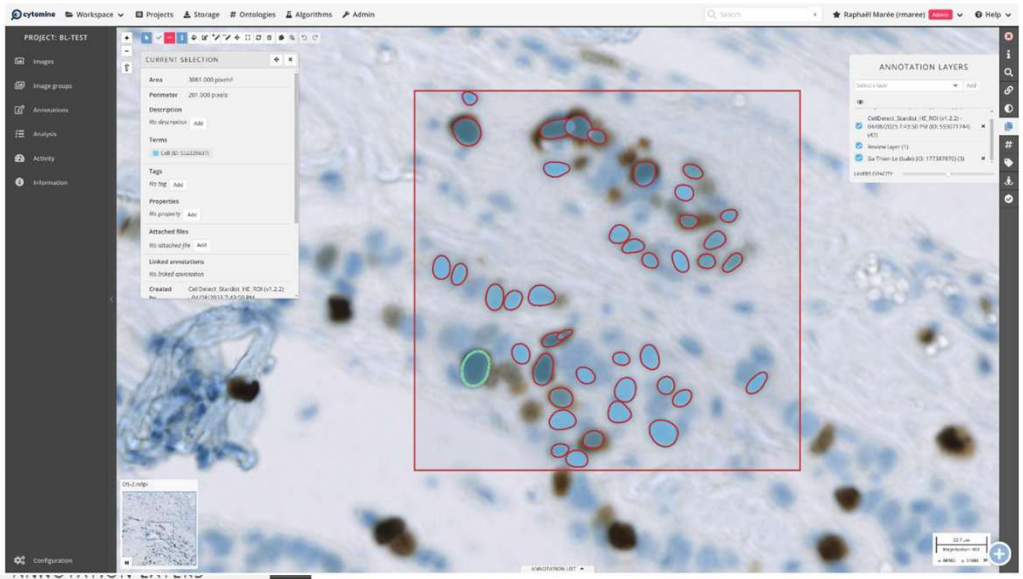
- 3 species
- 10 stains
- 13 human tissue types + toxicological tissues
- 13 scanners

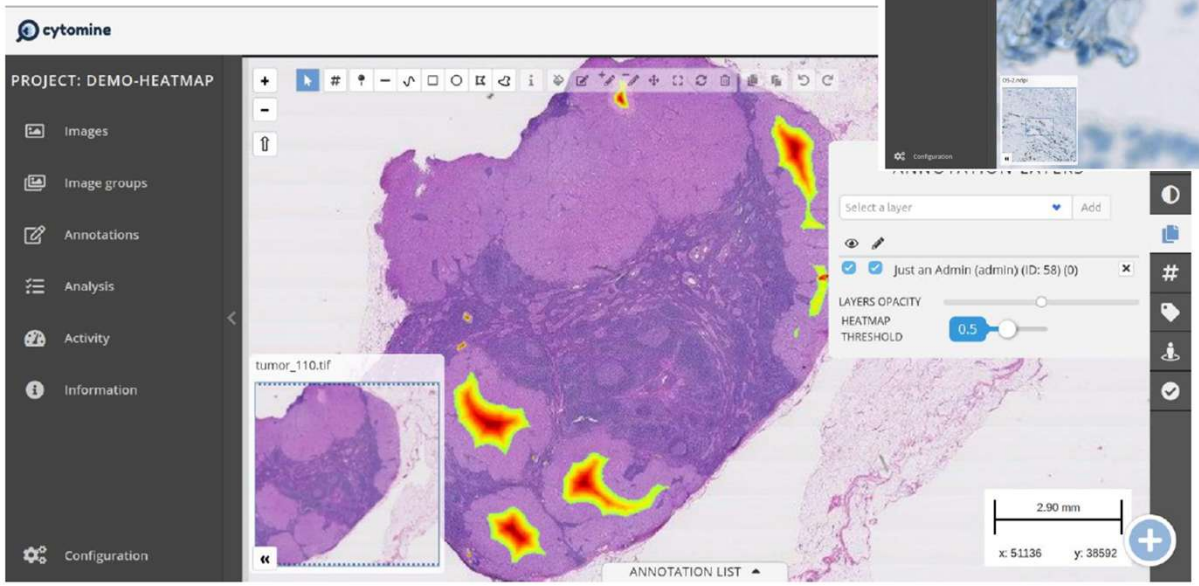
RESULTS

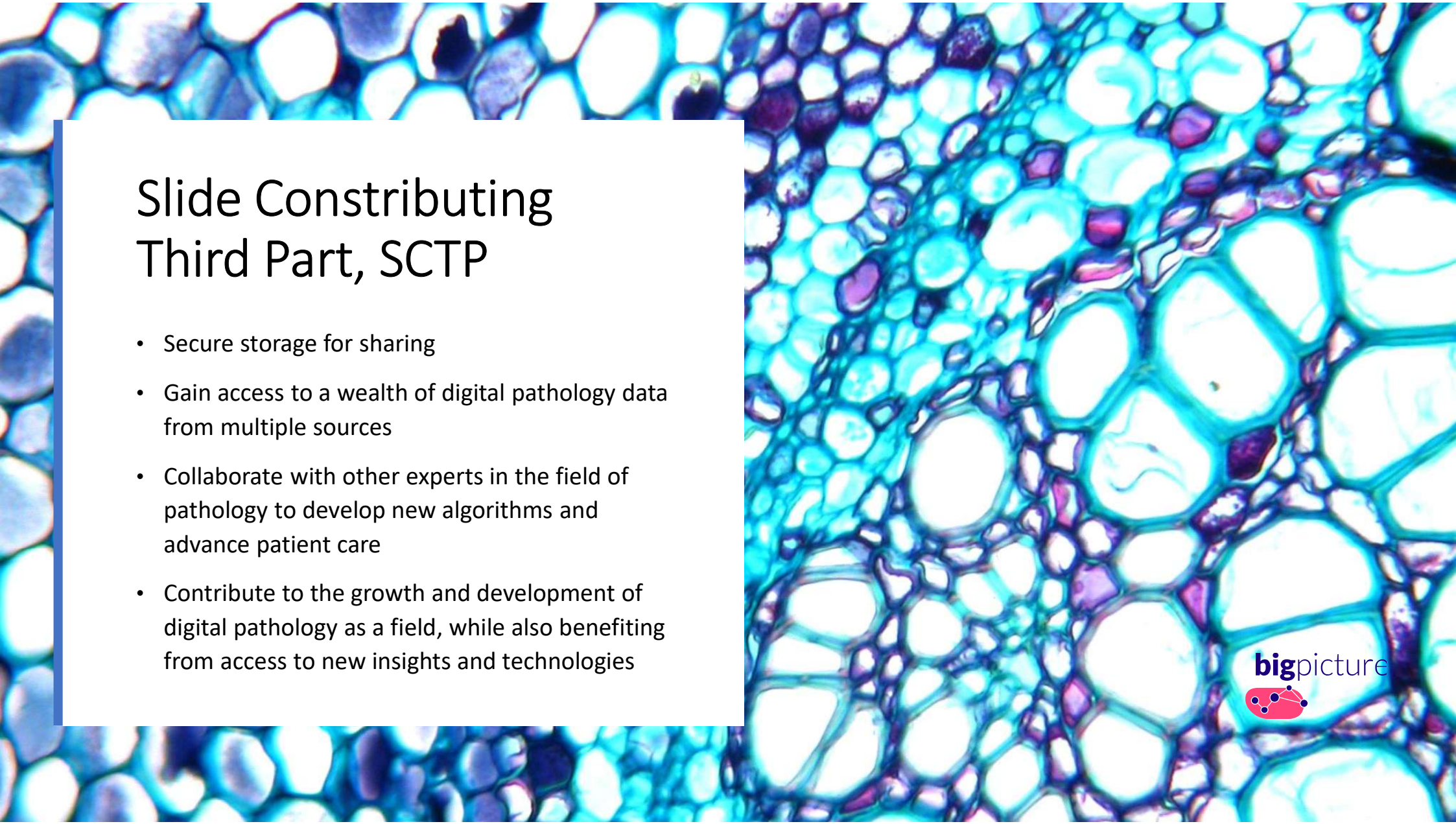
Artifact Class	Dice Score
Background	0.91
Tissue folds	0.68
Ink	0.82
Air bubbles	0.90
Dust	0.81
Markers	0.97
Out of focus	0.74
Average	0.92



AI-assisted annotation







Slide Constributing Third Part, SCTP

- Secure storage for sharing
- Gain access to a wealth of digital pathology data from multiple sources
- Collaborate with other experts in the field of pathology to develop new algorithms and advance patient care
- Contribute to the growth and development of digital pathology as a field, while also benefiting from access to new insights and technologies

bigpicture

Main Goals

- Accelerate the development of AI in pathology.
- Foster research, innovation and collaboration.

Align with Societal Needs

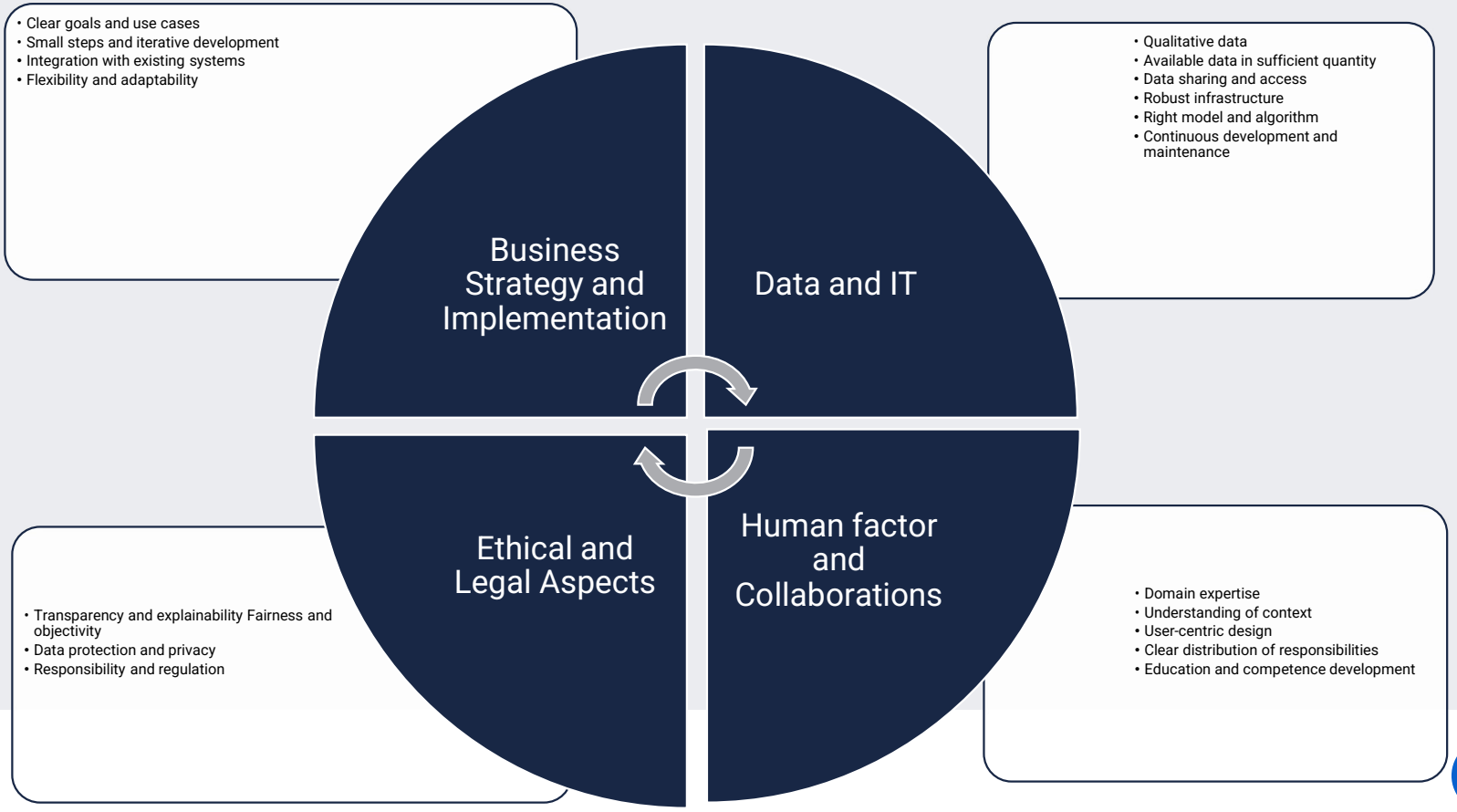
- Improve diagnostics, patient outcomes and cost-benefit.
- Support RD & innovation; foster community; promote computational pathology

Key Achievements

- User and contributors' needs and interests identified
- Products and services are defined
- Pricing and costs analysis was made
- Governance models have been developed
- Sustainable business models have been developed

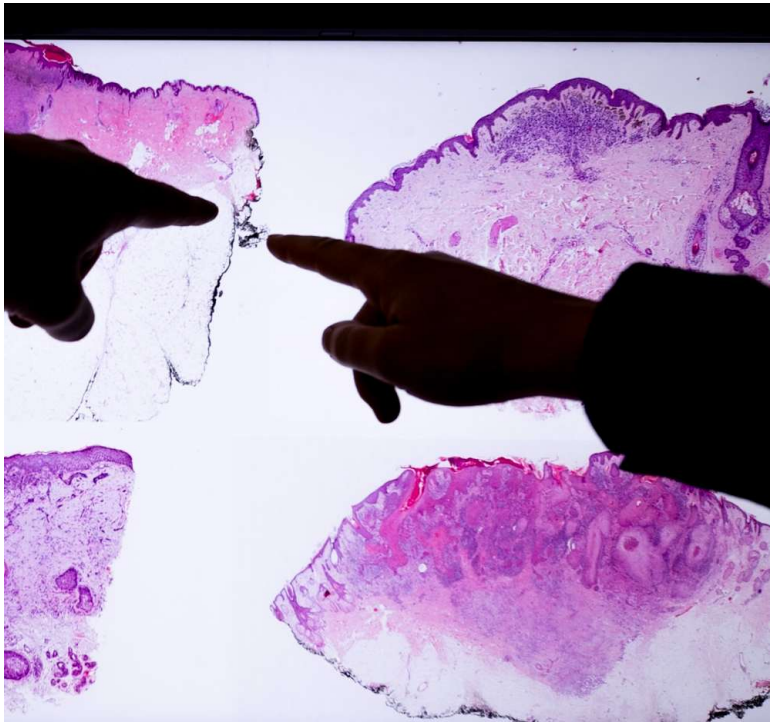
bigpicture

AI transformation



Region Östergötland

This project has received funding from the Innovative Medicines Initiative 2 Joint Undertaking under grant agreement No 945358. This Joint Undertaking receives support from the European Union's Horizon 2020 research and innovation program and EFPIA. imi.europe.eu



Thank you for your attention

bigpicture

Anna.C.Boden@regionostergotland.se

https://bigpicture.eu