OVERCOMING DATA SILOS:
THE AI COLLABORATION TOOLKIT FOR HEALTH DATA

Hans Meine

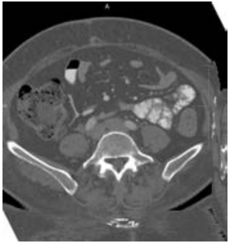
June 16th, 2021

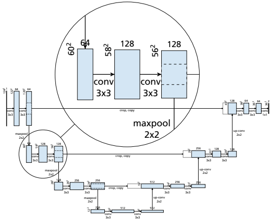
Research & Development of Data-Driven Algorithms

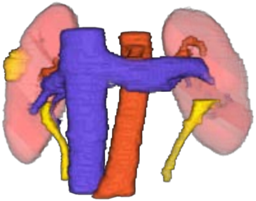
Example: Extraction of Kidney Anatomy

AI algorithm for assessment of renal cancer:

- Input: CA-CT (computer tomography) images of renal region
- Output: geometry of important anatomy (kidneys, tumors, vessels)





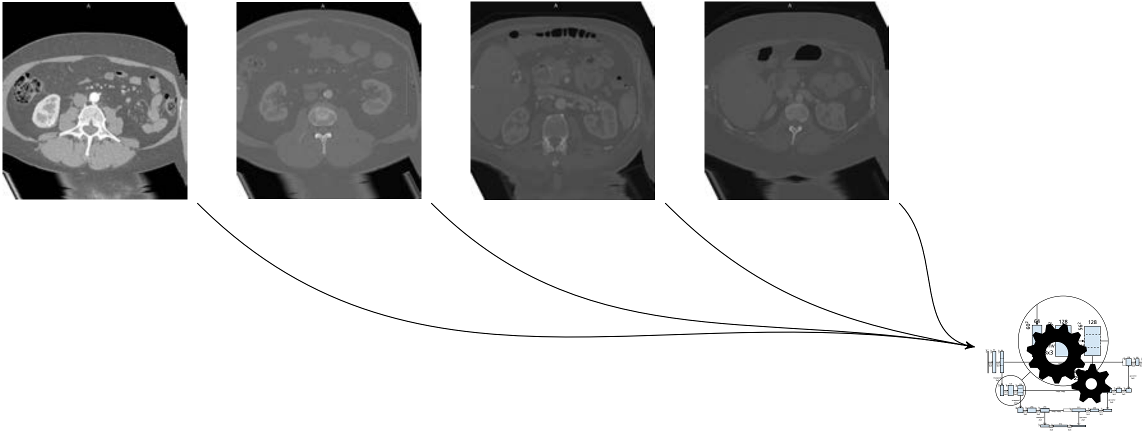


Goal: automatic measurements, quantitative assessment

Data used in the following: 2021 Kidney and Kidney Tumor Segmentation Challenge (KiTS 21)

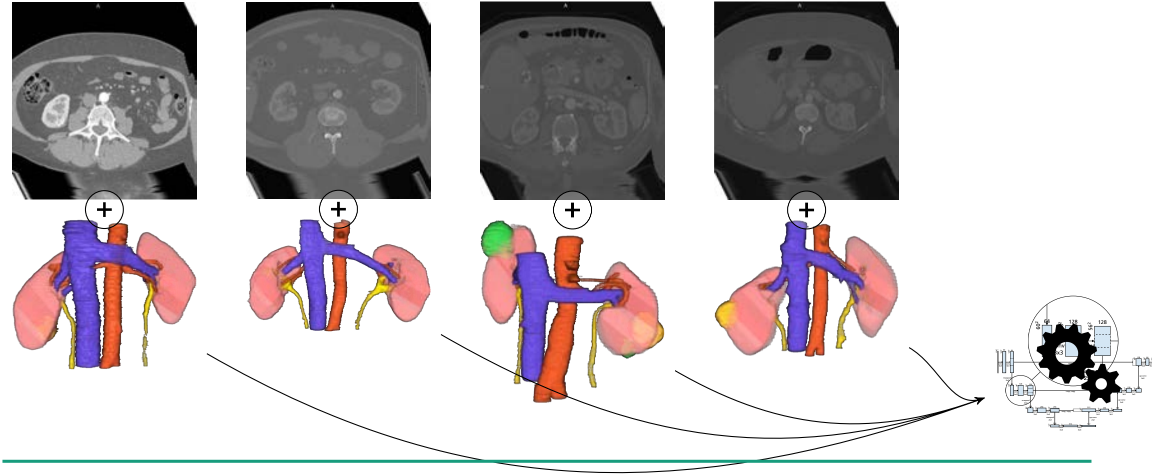
Pipeline for Training an AI (Supervised Learning)

Now how does the AI learn how to extract anatomy?



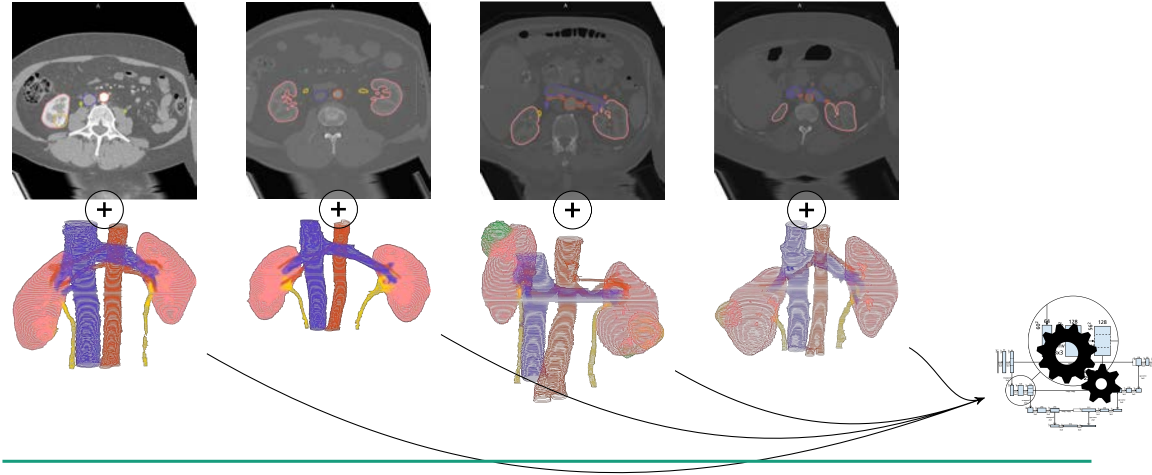
Pipeline for Training an AI (Supervised Learning)

Now how does the AI learn how to extract anatomy?



Pipeline for Training an AI (Supervised Learning)

Now how does the AI learn how to extract anatomy?

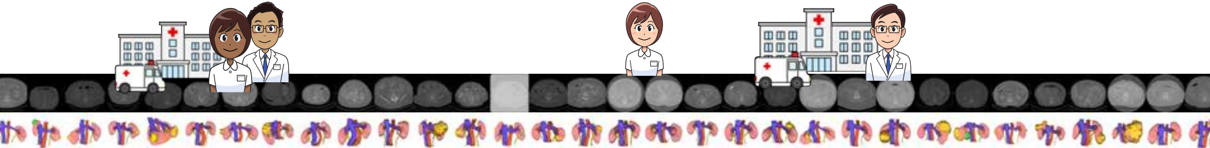


Complexities of Collaborative R&D

Sounds simple? In practice,

- data collected by doctors, needs to be accessible to technical scientists
  - anonymization has to be applied to data leaving the hospital
  - algorithm results or application prototypes have to be shared
- data needs to be reviewed / weeded out by both physicians and AI specialists
- annotations take a lot of time and need to be reviewed by experts
- data comes in batches, numbers increase over time
- problematic cases excluded at any time, annotations may need corrections, …

⇒ AI models are trained on *different versions* of curated data



PURPOSE / SCOPE

Our toolkit speeds up collaborative R&D of data-driven algorithms

Professional tooling for large, multi-site data collection & curation

Data curation

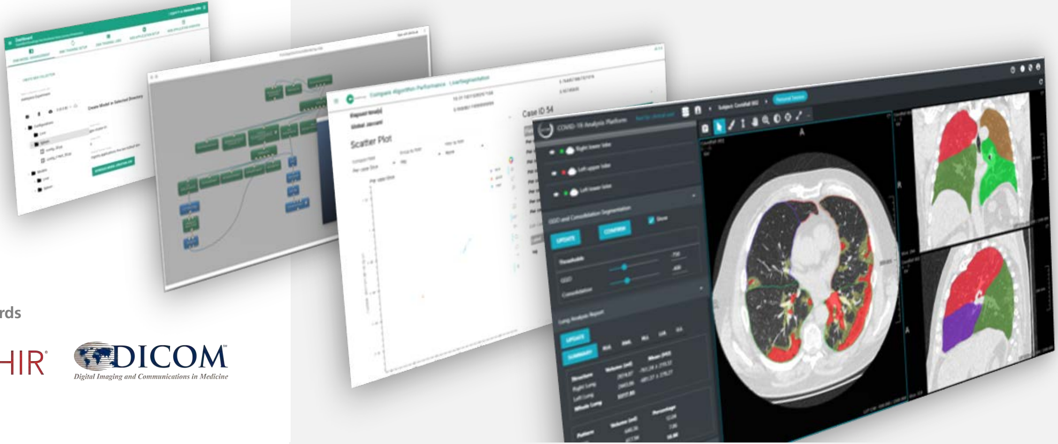
- Viewing
- Filtering
- Annotation

AI dev

- Model training
- Model evaluation

Application prototypes

- Clinical evaluation
- Workflow and UX assessment



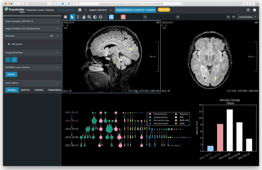
Based on standards

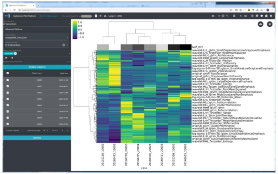
HL7 FHIR

DICOM
Digital Imaging and Communications in Medicine

SATORI – A Highly Customizable Annotation Tool

Main frontend visible to clinical users



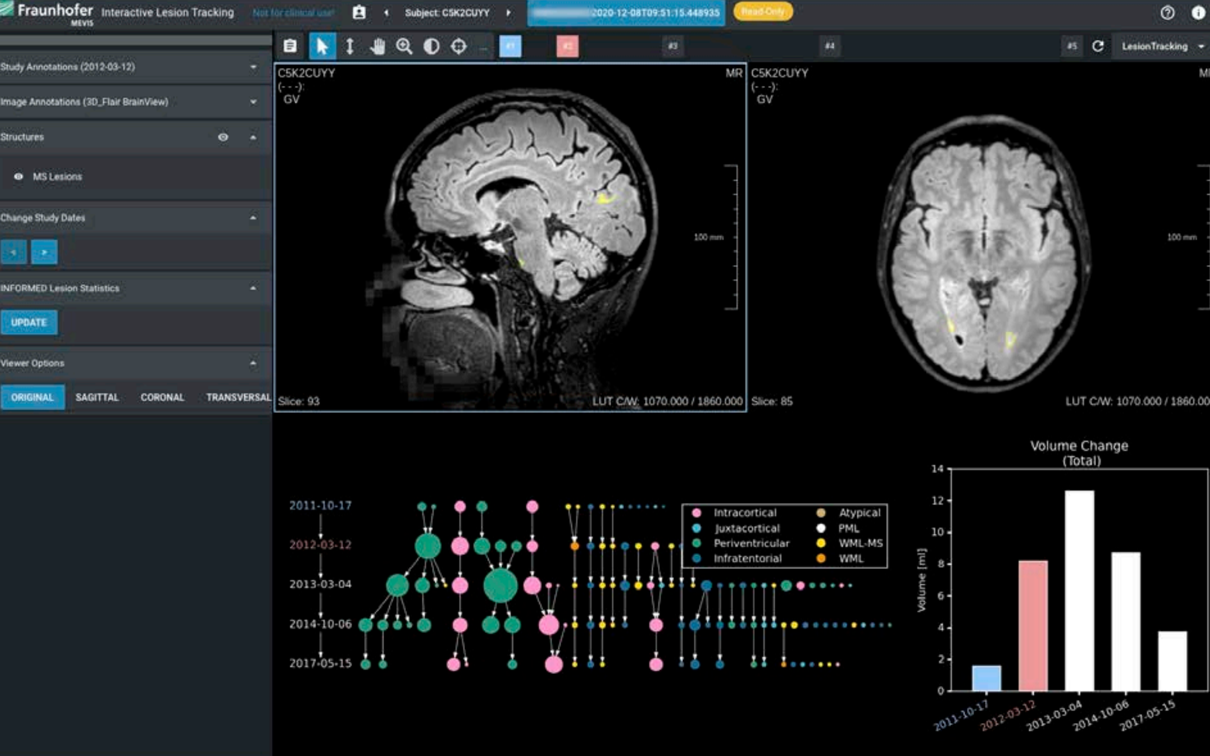


Basic Features

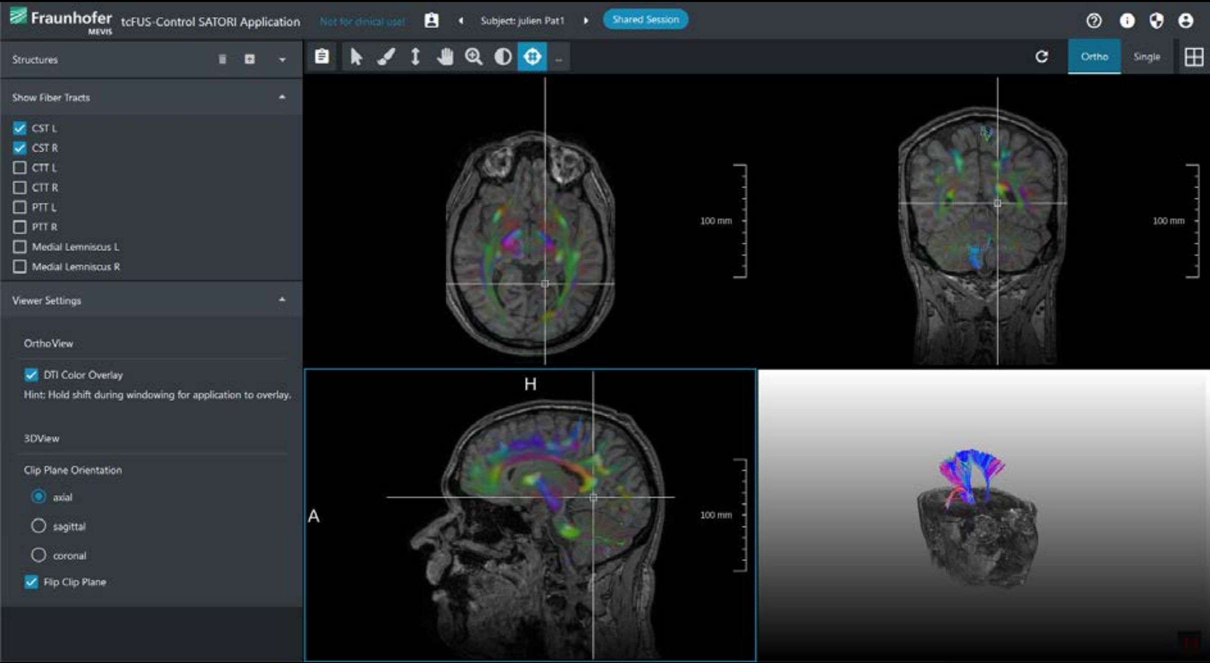
- Import / Export in many formats (DICOM, SEG, NIfTI, …)
- Automatic categorization, tagging and preprocessing
- User and group management, private / shared sessions, audit mode
- Customizable hangings and layouts
- Structured annotations (subject / study / image / structure)
- Efficient segmentation and correction of structures

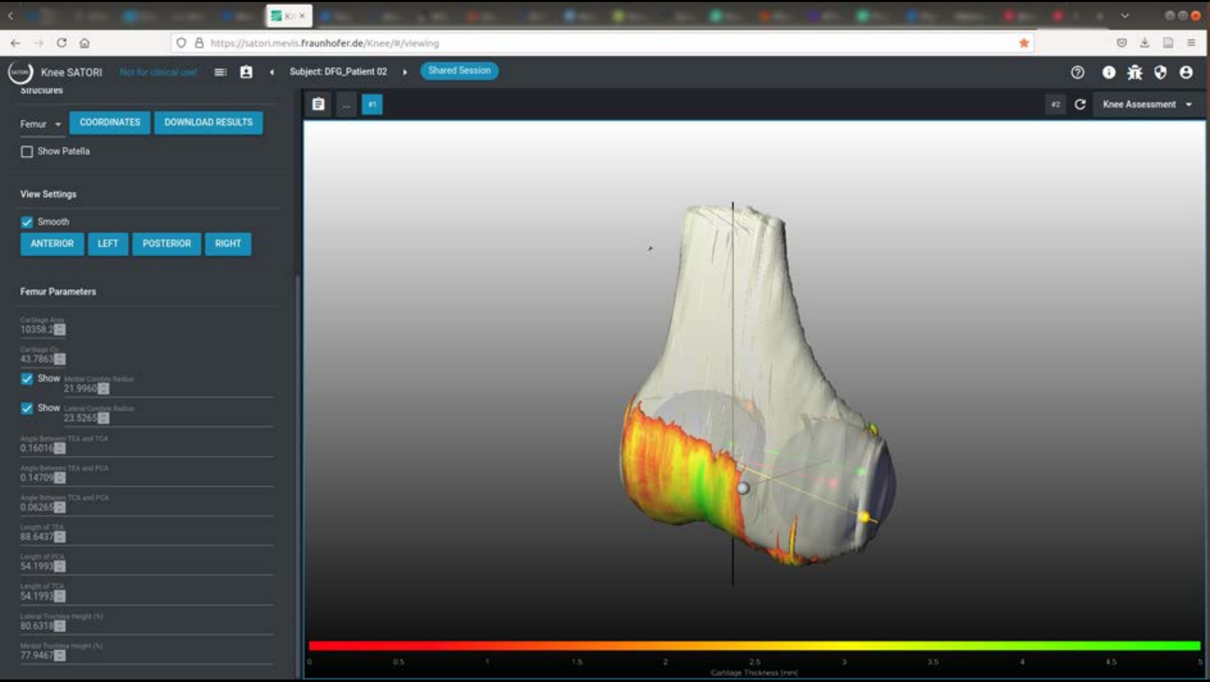
Advanced Features through extensibility

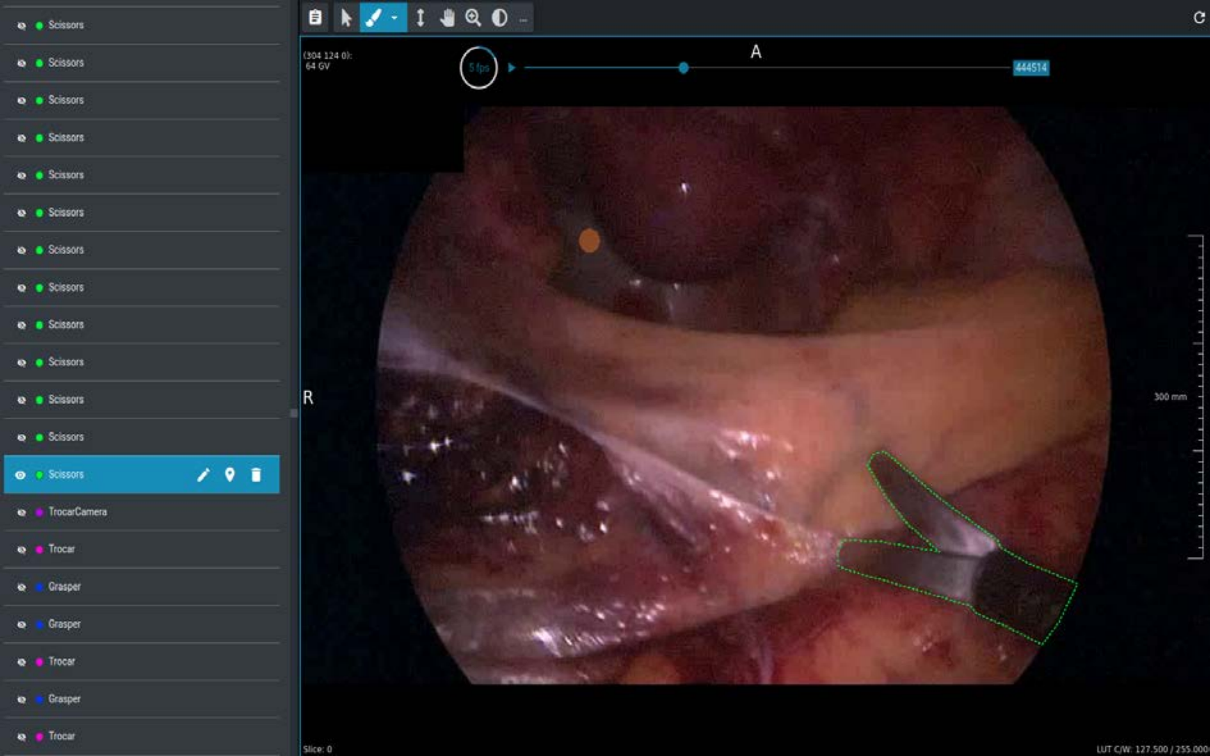
- Smooth transition towards application prototypes
- Deployment options (MEVIS, on site, cluster, cloud, …)
- Direct connection to Deep Learning



Fraunhofer
MEVIS
Interactive Lesion Tracking
Subject: C5K2CUYY
2020-12-08T09:51:15.448935
Study Annotations (2012-03-12)
Image Annotations (3D_Flair BrainView)
Structures
MS Lesions
Change Study Dates
INFORMED Lesion Statistics
UPDATE
Viewer Options
ORIGINAL
SAGITTAL
CORONAL
TRANSVERSAL
#3
#4
#5
LesionTracking
C5K2CUYY
GV
MR
Slice: 93
LUT C/W: 1070.000 / 1860.000
Slice: 85
LUT C/W: 1070.000 / 1860.00
100 mm
2011-10-17
2012-03-12
2013-03-04
2014-10-06
2017-05-15
Intracortical
Juxtacortical
Periventricular
Infratentorial
Atypical
PML
WML-MS
WML
Volume Change
(Total)
Volume [ml]
0
2
4
6
8
10
12
14


Fraunhofer
MEVIS
tcFUS-Control SATORI Application
Not for clinical use!
Subject: julien Pat1
Shared Session
Structures
Ortho
Single
Show Fiber Tracts
CST L
CST R
CTT L
CTT R
PTT L
PTT R
Medial Lemniscus L
Medial Lemniscus R
Viewer Settings
OrthoView
DTI Color Overlay
Hint: Hold shift during windowing for application to overlay.
3DView
Clip Plane Orientation
axial
sagittal
coronal
Flip Clip Plane
100 mm
100 mm
H
A
100 mm


https://satori.mevis.fraunhofer.de/Knee/#/viewing
Knee SATORI
Not for clinical use!
Subject: DFG_Patient 02
Shared Session
Knee Assessment
Structures
Femur
COORDINATES
DOWNLOAD RESULTS
Show Patella
View Settings
Smooth
ANTERIOR
LEFT
POSTERIOR
RIGHT
Femur Parameters
10358.2
43.7863
Show
21.9960
Show
23.5265
0.16016
0.14709
0.06265
88.6437
54.1993
54.1993
80.6318
77.9467
0
0.5
1
1.5
2
2.5
3
3.5
4
4.5
5


Scissors
Scissors
Scissors
Scissors
Scissors
Scissors
Scissors
Scissors
Scissors
Scissors
Scissors
Scissors
Scissors
TrocarCamera
Trocar
Grasper
Grasper
Trocar
Grasper
Trocar
(304 124 0):
64 GV
5 fps
444514
A
R
300 mm
Slice: 0
LUT C/W: 127.500 / 255.000

Radiomics Pilot Platform
localhost:8088/#/radiomicsAnalysis
Radiomics Pilot Platform
Not for clinical use!
You are logged in as
Subject: 10001
LV Epicardium
Advanced Options
exampleMR_5mm.yaml
LV Endocardium
LVEF-MRT
RETRIEVE SUBJECTS
Subject name
Study time
10002
2013-11-18
10002
2016-04-22
10001
2013-12-09
10001
2014-06-03
20003
2016-04-11
20003
2016-07-18
20004
2016-07-05
20004
2016-09-20
8 records selected.
Records per page:
10
1-8 of 8
ANALYZE
1.6
0.8
0.0
−0.8
−1.6
lvef_mrt
wavelet-LLL_gldm_SmallDependenceLowGrayLevelEmphasis
wavelet-LHL_firstorder_RootMeanSquared
wavelet-HLH_glrlm_RunVariance
wavelet-HHH_glszm_LargeAreaEmphasis
wavelet-LLH_firstorder_Median
wavelet-LHH_firstorder_Uniformity
wavelet-LHH_glcm_InverseVariance
log-sigma-5-0-mm-3D_glszm_SmallAreaLowGrayLevelEmphasis
wavelet-LHL_glszm_ZoneVariance
original_glrlm_RunVariance
original_gldm_GrayLevelNonUniformity
log-sigma-3-0-mm-3D_glcm_InverseVariance
wavelet-HLL_gldm_LargeDependenceEmphasis
wavelet-LHH_firstorder_Energy
wavelet-HLL_glszm_LowGrayLevelZoneEmphasis
wavelet-LHH_firstorder_RootMeanSquared
wavelet-HHL_glszm_SmallAreaLowGrayLevelEmphasis
wavelet-LLH_glrlm_HighGrayLevelRunEmphasis
log-sigma-3-0-mm-3D_firstorder_Mean
wavelet-LHL_firstorder_Range
wavelet-HLL_glcm_Autocorrelation
wavelet-HLL_glcm_ClusterProminence
wavelet-HHH_glcm_Idmn
wavelet-HHH_glcm_Autocorrelation
wavelet-LHH_firstorder_Range
wavelet-LLL_glcm_JointAverage
wavelet-HLH_firstorder_RobustMeanAbsoluteDeviation
wavelet-LHH_firstorder_MeanAbsoluteDeviation
wavelet-HHH_firstorder_90Percentile
original_glcm_ClusterProminence
wavelet-LLH_gldm_GrayLevelVariance
wavelet-LHH_gldm_DependenceEntropy
log-sigma-3-0-mm-3D_gldm_SmallDependenceEmphasis
wavelet-LLH_glrlm_RunPercentage
original_glrlm_RunLengthNonUniformityNormalized
wavelet-HHL_firstorder_Entropy
20131118_10002
20160422_10002
20160718_20003
20160705_20004
20160920_20004
20160411_20003
20131209_10001
20140603_10001
case

Technology Goal 1: In-Situ Correction

Only dynamically learning AI appears really "intelligent"

Long history of application prototypes with manual correction facilities

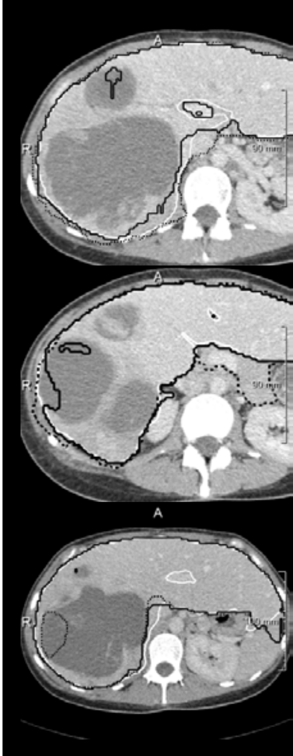
- Doctors become frustrated when software repeats the same mistake again and again

Our highly integrated web-based AI collaboration toolkit makes these corrections more sustainable

- Corrections enter data pool…
- …undergo review / curation…
- …and can be used for training / evaluation.

Thought on Deep Learning

We expect corrections to be *particularly valuable*!

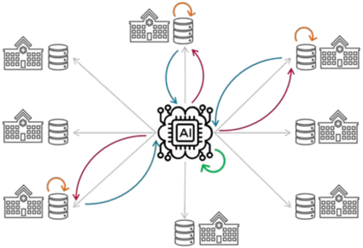


Technology Goal 2: Federated Learning

Because multi-centric data is key to robust AI & real-world applicability

Toolkit supports on-site deployment

⇒ **data does not have to leave the hospital**



Trained models may travel, though

- for evaluation / application
- for learning on data from multiple sites

Our toolkit provides interfaces also for models

- goal: supporting real-time communication betw. hospitals

Problem dimensions

- hurdles for data sharing
- firewalled IT (as a consequence)
- training strategies

Conclusions

We are building an AI Collaboration Toolkit with integrated platform components for

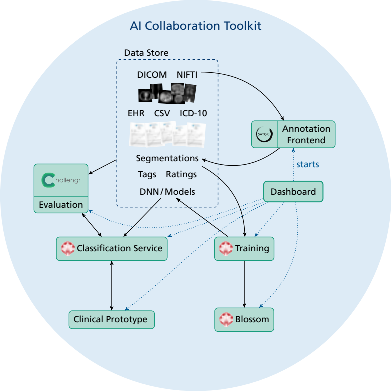
- data curation
- AI research & development
- custom application prototypes

Integration enables

- dynamically learning AI and
- federated learning

The toolkit is used in ongoing projects, internally and with various partners

- RACOON-SATORI currently deployed to all German UMC (NUM)



Fraunhofer MEVIS