

Pathology Whole Slide Image De-Identification

MIDI Workshop
Day 2



SageBionetworks

Panelists

Adam Taylor



David Gutman



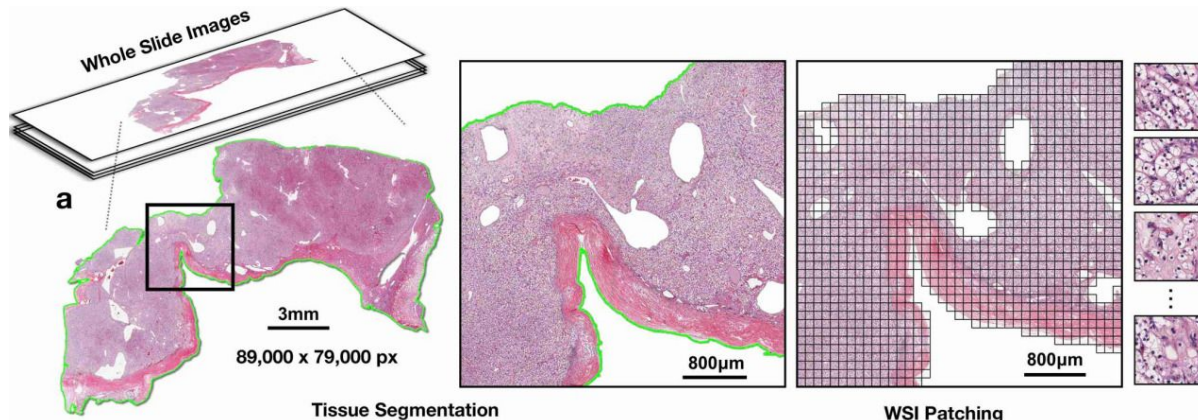
Tom Bisson



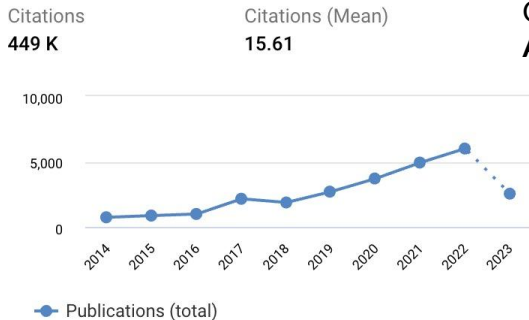
Format

- Introduction
 - Digital pathology and whole slide imaging in the context of de-identification
 - De-identification validation for WSI and multiplexed tissue imaging in HTAN
- Presentations on recent work
 - David Gutman
 - Tom Bisson
- Panel discussion

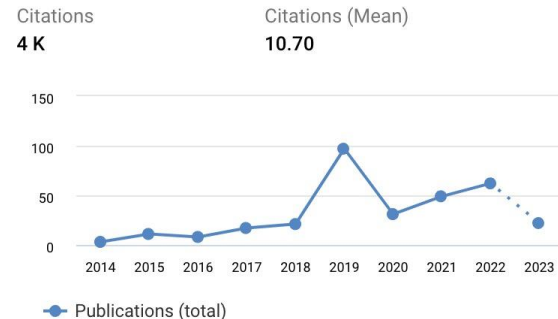
Digital pathology and whole slide imaging



("digital pathology"
OR "whole slide imaging")



("digital pathology"
OR "whole slide imaging")
AND "de-identification"



De-identification considerations for WSI

Filename

TCGA-V1-A8MU-01Z-00-DX1.2C9CED13-5C2C-4FFB-AB0B-3904BAA4FEFF.svs (864,7 MB)

Slide label



Burnt in text



Image headers/tags

Assay/study metadata

TIFF public tags

DateTime 306 (0x132) ASCII: 2020:01:28 14:32:26

Format and metadata diversity

Bioformats supports 162 formats with varying degrees of metadata readability

Amersham Biosciences Gel	.gel	▲	▲	▲	▼	▼	✖	✖	✖	✖
Amira Mesh	.am, .amiramesh, .grey,.hxx, .labels	▲	▼	▼	▼	▼	✖	✖	✖	✖
Amnis FlowSight	.cif	▲	▲	▲	▼	▼	✖	✓	✓	✖
Analyze 7.5	.img,.hdr	▲	▲	▲	▲	▼	✖	✖	✖	✖
Andor SIF	.sif	▲	▼	▼	▼	▼	✖	✖	✖	✖
Animated PNG	.png	▲	▲	▲	▲	▼	✓	✓	✖	✖
Aperio AFI	.afi,.svs	▲	▲	▲	▲	▲	✖	✖	✓	✓
Aperio SVS TIFF	.svs	▲	▲	▲	▲	▲	✖	✖	✓	✓
Applied Precision CellWoX	.htd,.pnl	▲	▲	▲	▼	▼	✖	✖	✓	✖
AVI (Audio Video Interleave)	.avi	▲	▼	▼	▲	▼	✓	✓	✖	✖
Axon Raw Format	.arf	▲	▲	▲	▼	▼	✖	✖	✖	✖
BD Pathway	.exp,.tif	▲	▲	▲	▼	▼	✖	✖	✓	✖
Becker & Hickl SPC FIFO	.spc	▼	▼	▲	▼	▼	✖	✖	✓	✖
Becker & Hickl SPCImage	.sdt	▲	▲	▲	▼	▼	✖	✖	✓	✖

TCGA-V1-A8MU-01Z-00-DX1.2C9CED13-5C2C-4FFB-AB0B-3904BAA4FEFF.svs (864,7 MB)

debuggingtissue.com

Level: 1 (baseline)
Resolution: 81671x64173
Downsample factor: 1

L: 2
R: 20417x16043
DF: 4

L: 3
R: 5104x4010
DF: 16

L: 4
R: 2552x2005
DF: 32

Associated images

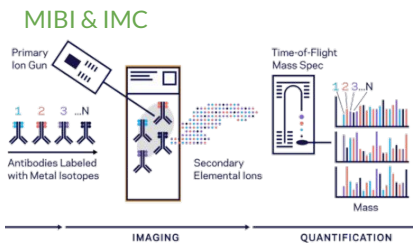
Label
R: 663x652

Macro
R: 1600x631

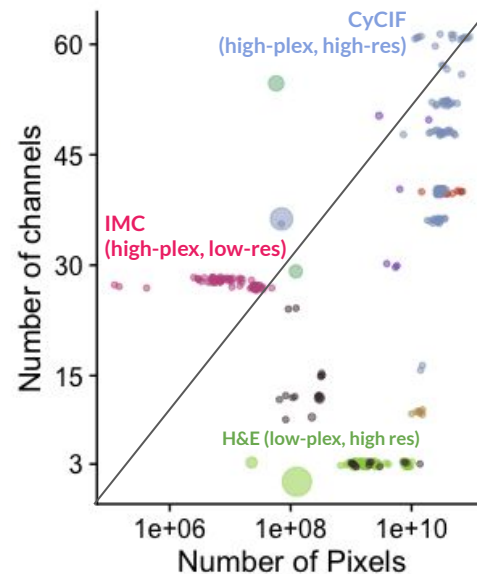
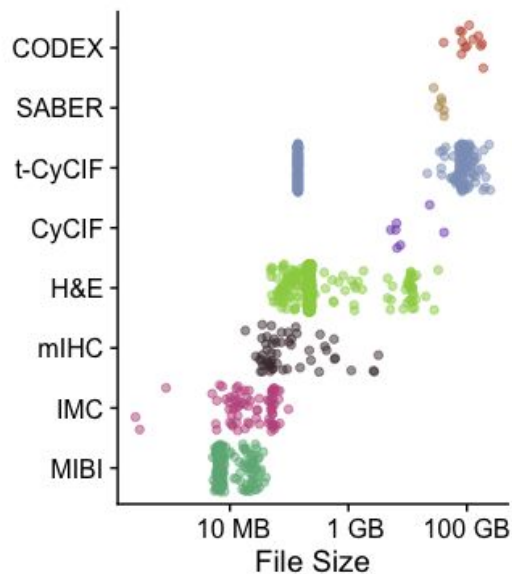
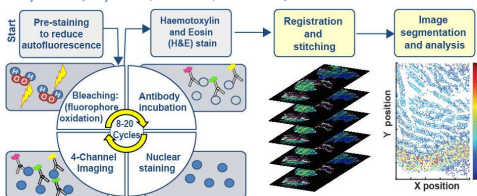
Thumbnail
R: 977x768

Multiplexed tissue imaging

We are in an expansionist phase in the development of advanced multiplexed tissue imaging modalities. We should anticipate their adoption in research and clinical environments



Cyclic (CyCIF, MxIF, mIHC)



Increasingly larger, higher dimension and higher resolution data

Sage Bionetworks

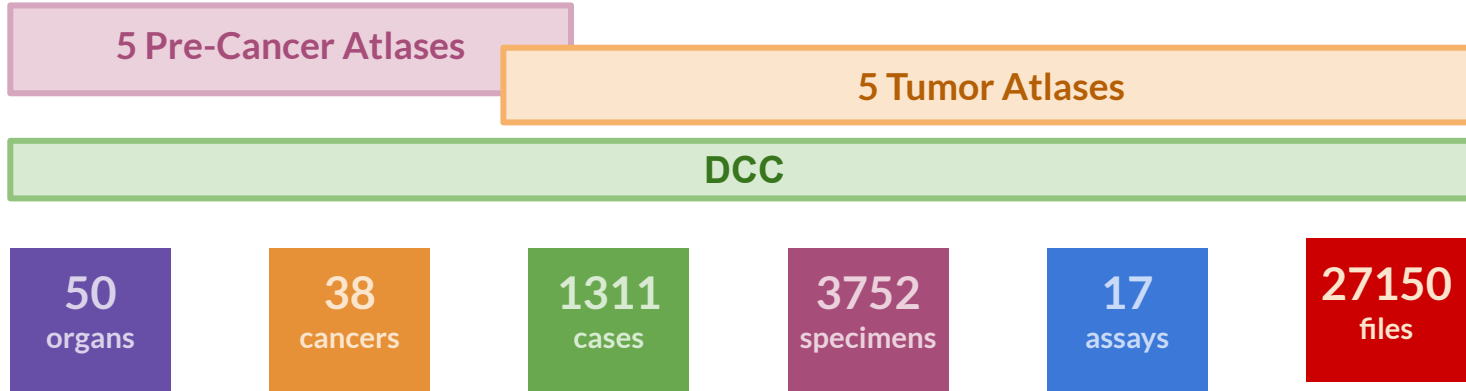
Accelerate biomedical discoveries by improving methods for scientific *collaboration* and *communication*



Human Tumor Atlas Network

A National Cancer Institute funded Cancer Moonshot Initiative

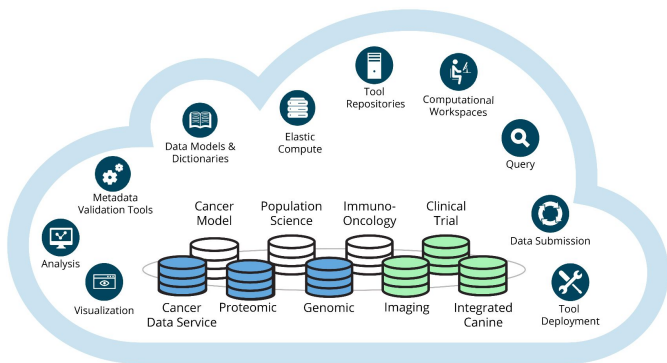
Normal → Pre-malignant → Malignant → Metastatic → Responsive → Resistant



Project goal(s): Construct 3D atlases of the dynamic cellular, morphological and molecular features of human cancers as they evolve from precancerous lesions to advanced disease.

The Human Tumor Atlas Network: Charting Tumor Transitions across Space and Time at Single-Cell Resolution.
Cell. 2020. doi:10.1016/j.cell.2020.03.053

HTAN Data Destinations



CANCER GENOMICS CLOUD
SEVEN BRIDGES

A RESOURCE OF THE NCI CANCER RESEARCH DATA COMMONS

ISB-CGC

Cancer Gateway in the Cloud

Access, Explore and Analyze Large-Scale Cancer Data Through the Google Cloud



Google
Big Query

Level 1 and 2 Sequencing →

NIH CRDC CDS
authenticated through dbGaP
available through Seven Bridges
Genomics Cloud

Level 3 and 4 Sequencing →

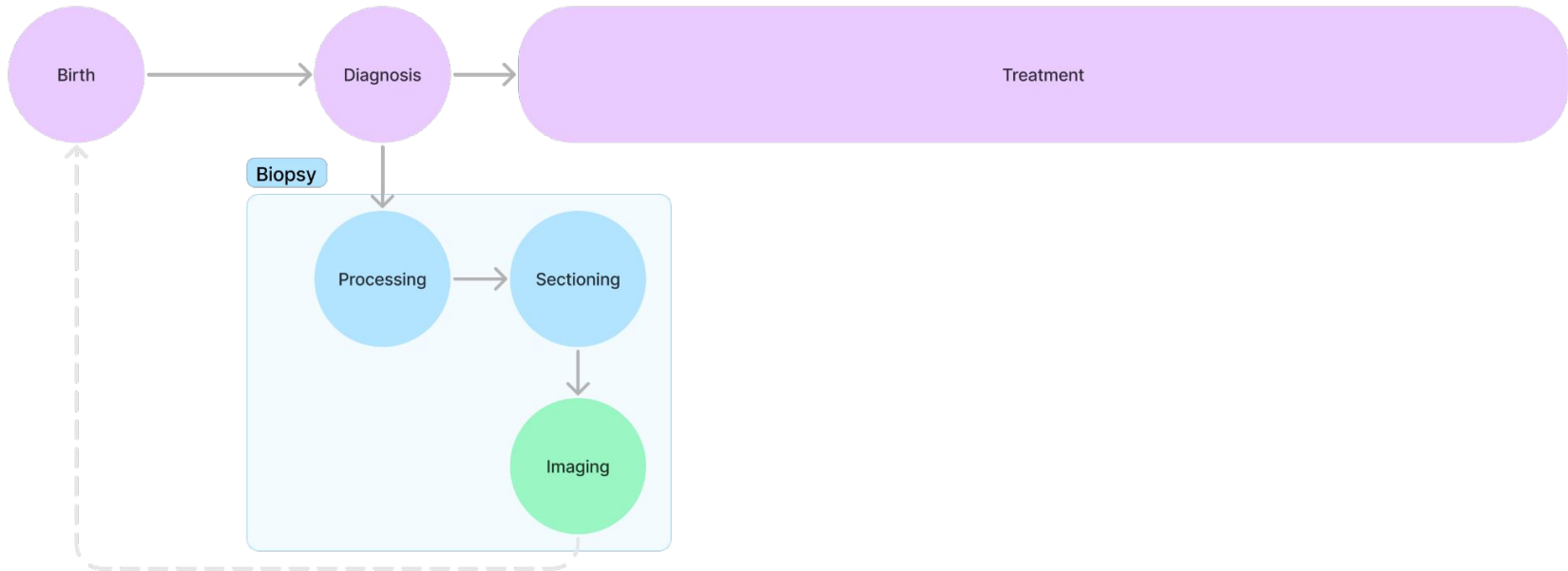
Available through Synapse, and
through Google BigQuery at
ISB-CGC (June 2022)

Imaging →

CDS and Imaging Data Commons

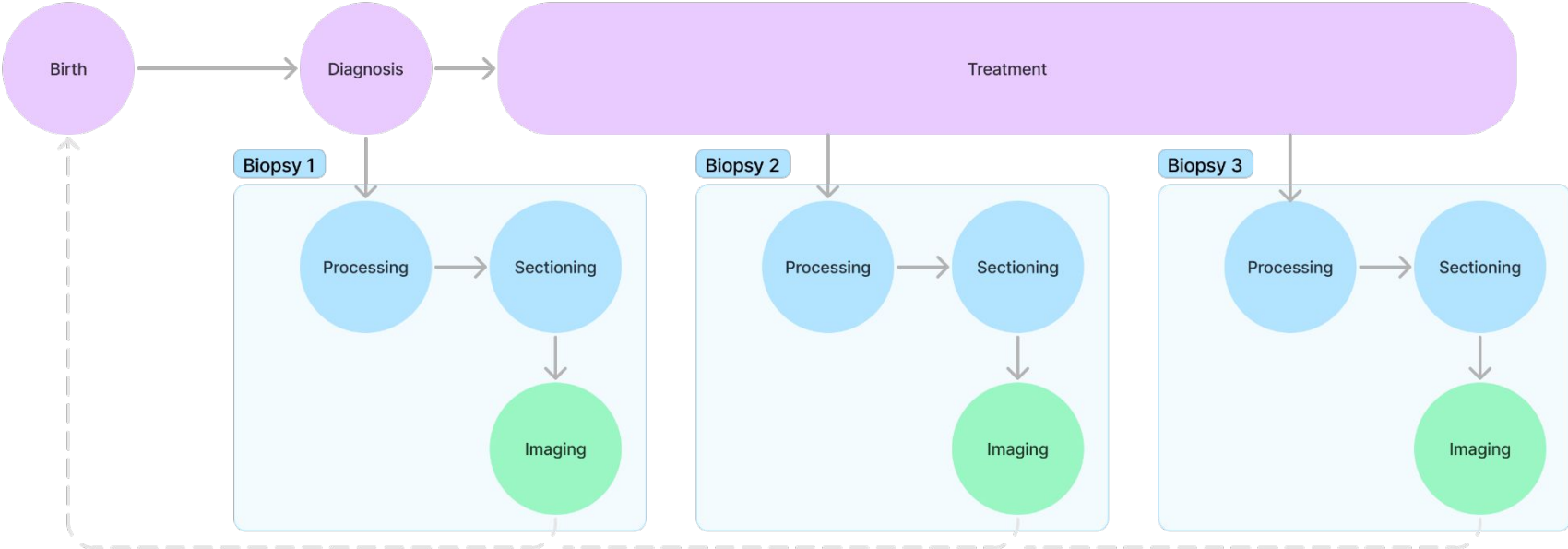
Data in the context of additional metadata

Metadata such as 'Sectioning days to index' where sectioning is proximal to imaging provides an attack vector to reconstruct participant date of birth

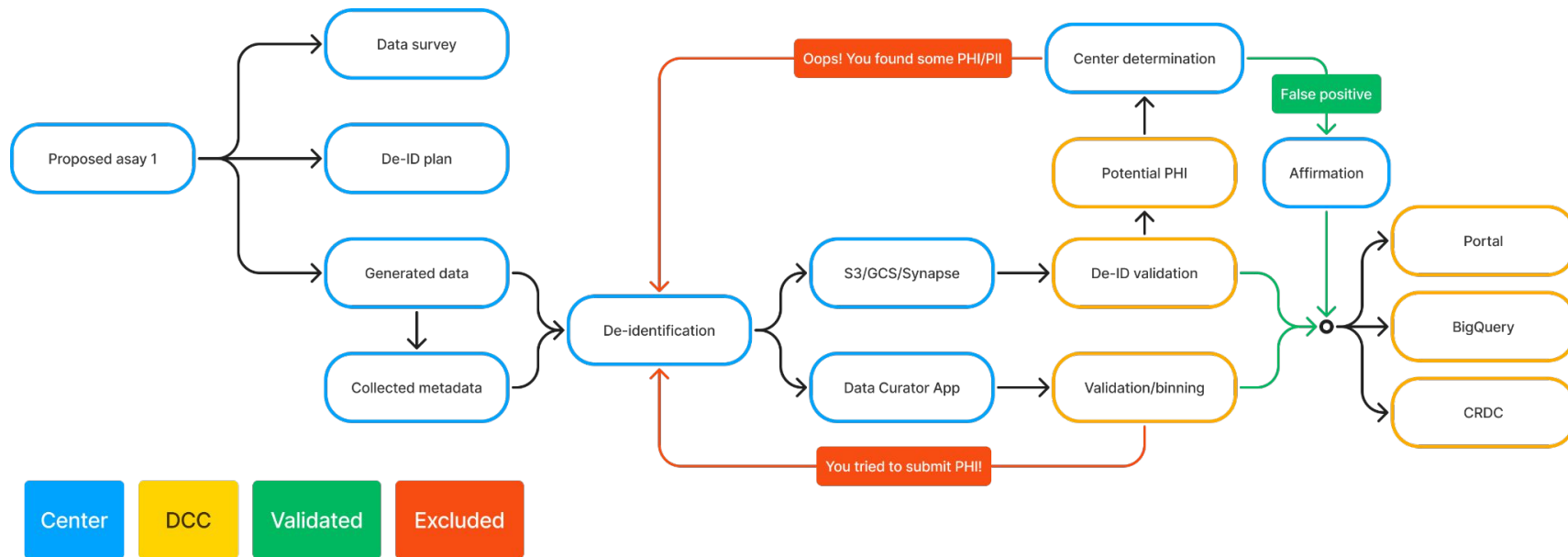


Data in the context of additional metadata

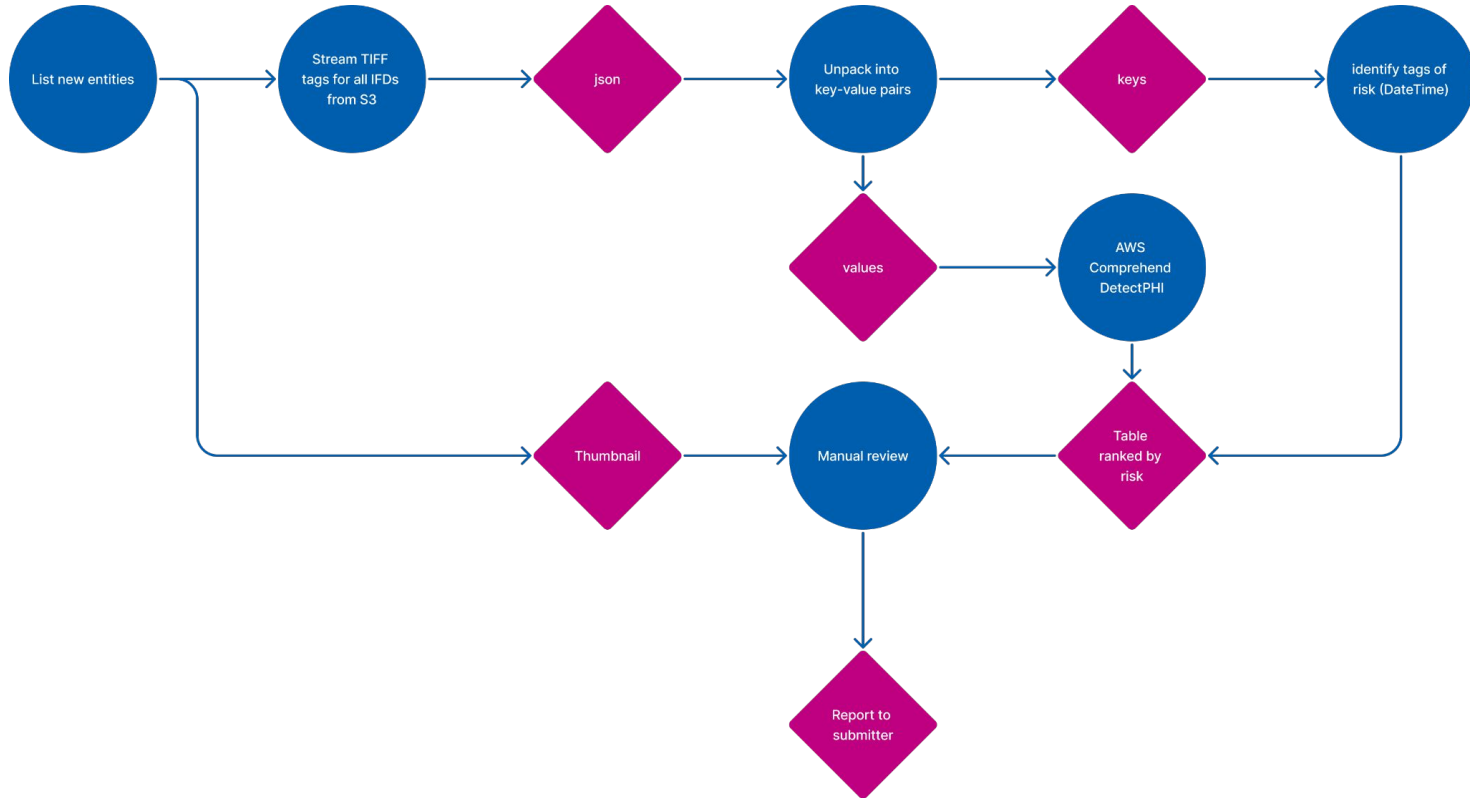
This effect is multiplied in the case of longitudinal studies



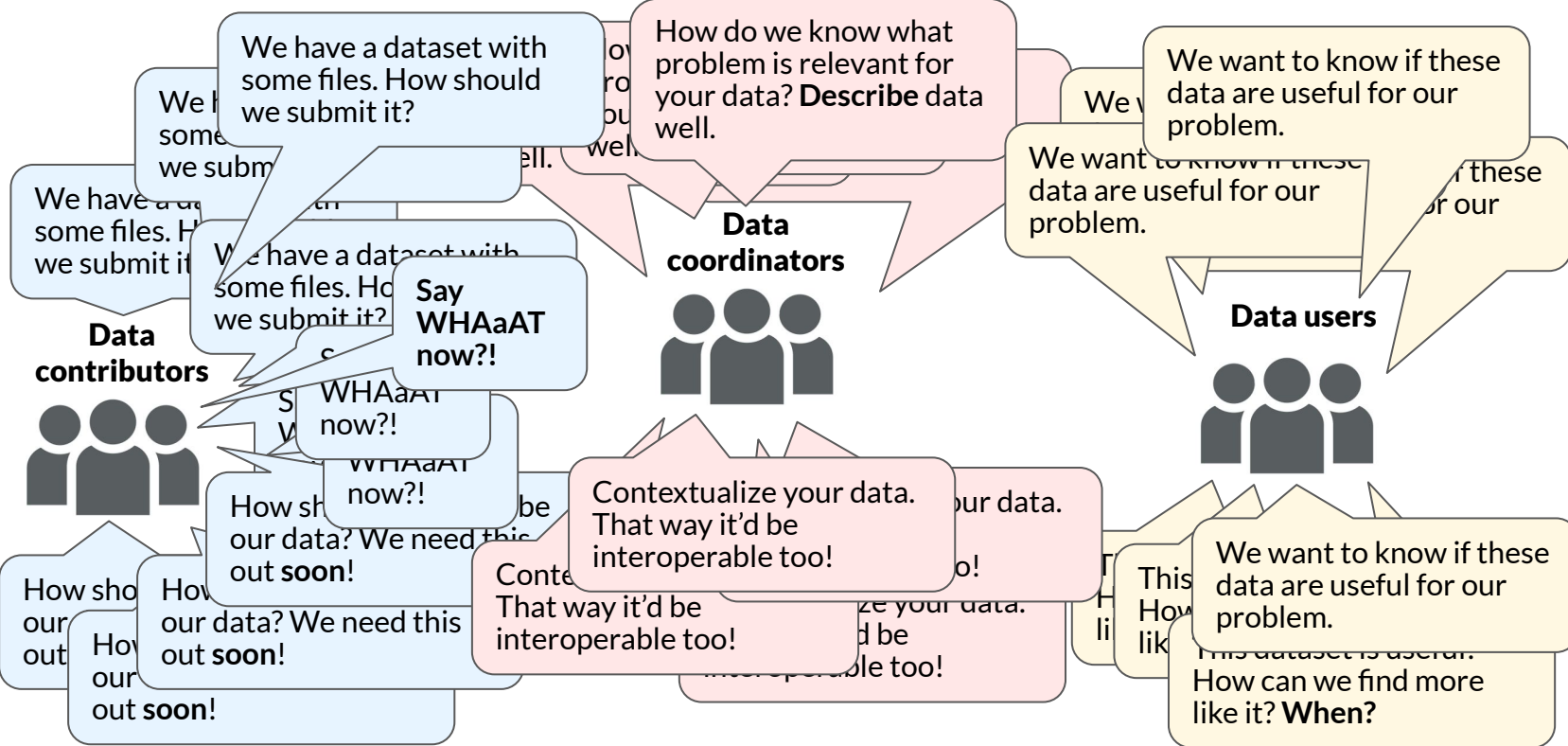
Validation of de-identification in HTAN



Validation of de-identification in HTAN



FAIR data requires communication at scale



Adam Taylor



Sage Bionetworks

David Gutman



EMORY
UNIVERSITY

Tom Bisson

