2021 MIDI Task Group Meeting Notes

July 20, 2021 Meeting

WebEx recording of 7/20/2021 meeting

- Introduction: Medical Image De-Identification Initiative (MIDI)
- Task Group goals
- Steering Committee
- Timeline
- Discussion

August 10, 2021 Meeting

WebEx recording of 8/10/2021 meeting

- Instructions to access the MIDI Task Group wiki page
- Accept Mendeley invitation to access private group for literature review/annotated bibliography
- Outline of approach
    - metadata vs. pixel data
    - structured (strongly typed) vs. text
    - burned-in text ("printed" and hand-written)
    - identifiable features (e.g., faces, iris, retina)
    - with or without "public" data to compare with
- Challenging topics
    - evaluation of success of de-identification
    - quantitative comparison of performance
    - quantifying re-identification risk
    - creating test data sets
    - faces (etc.) reconstructed from cross-sections
    - burned-in text - detection, removal, cleaning
    - cleaning text descriptors (metadata or burned in)
    - buried metadata (e.g., EXIF, geotags in JPEG inside DICOM)
    - dates (incl. preserving temporal relationships)
    - pseudonym consistency across separate submissions
    - risks of hashing to create pseudonymous identifiers
    - uniqueness of images limits statistical approaches
    - loss allowable during de-identification (e.g., age fuzzing, pixels)
    - private data element preservation to retain utility
    - ultrasound - still frames and cine loops, lossy compressed
    - photographs and video
    - gross pathology and whole slide images (incl. labels)
    - IRB/ethics committee messaging wrt. de-identification decisions
    - IT security approval/audits of de-identification
    - regulatory requirements: HIPAA Privacy Rule, GDPR, CCPA, others?
    - sufficiency of standards, e.g., DICOM PS3.15 Annex E
    - risk of not following a standard (home-grown decisions)
    - threat of image "signatures", private set intersection methods
    - policy versus the technical details of recompression/decompression artifacts for JPEG
    - data minimization
- Inventory of tools
    - user interface vs. scripted (bulk, service)
    - configurable - user vs. installer vs. hard-coded
    - platform, language
    - open source, free, commercial, service
    - on-site vs. outside (e.g., [IP]II needs to leave walls for AI on cloud)
- Roadmap and deliverables
    - interim report
        - full report
        - "primer" on medical image de-identification for newbies/execs
        - confirm what is out of scope (non-goals) - consent, data use agreements, ...
- Tasking: Members to think about which task they would like to contribute to.

September 14, 2021 Meeting

WebEx recording of the 9/14/2021 meeting

- Role of AI in de-identification - demand for data, opportunities, threats
    - Google has a de-id tool
    - Amazon Comprehension
    - Identifying images at risk–which images are likely to contain burned in information than others?
    - Problem with scalability in terms of building the ruleset. Better to identify selectively.
    - Barcodes, pacemaker serial numbers, implanted devices
    - There is the potential of identifying objects but not the raw data.
    - Action: Describe the steps involved in imaging and the evolution of data in different levels of processing
      Case-based data
    - Is raw data in our purview?
    - Raw data is often in proprietary format and can lack a header.
    - Post-processed data like 3D reconstructions
    - What is the harm of reidentification? High-resolution 3D image of the face
    - Penetration testers that applies to de-ID
    - How to evaluate the success of de-facing?
    - Newman, L. H. (2016). AI Can Recognize Your Face Even If You're Pixelated. Wired. https://www.wired.com/2016/09/machine-learning-can-identify-pixelated-faces-researchers-show/
    - When is it okay to release information that you know is identifiable? Example of boy in NYT.
    - Sometimes reidentification does not provide any new data.
    - What do you now know that you didn't know before?
    - Expectations of doing better deidentification and the threats of better reidentification. What can we do now and what in the future with AI?
    - Do you expect that one day a machine will replace your manual deidentification process? Can a robot replace human review?
    - Can you accept the risk of AI/machines/code? Get to the level of risk that is tolerable.
    - Main topic for the next call: the need for human QC.
    - When will you stop using humans or a targeted subset?
    - What would increase your comfort level to help you stop using human QC.

October 12, 2021 Meeting

WebEx recording of the 10/12/2021 meeting

Discussion of this document:

- Not practical for a human to review all of the images.
- TCIA built a tool called Kaleidoscope that flattens images and saves time.
- Radiology techs can also do this work, but sensitivity goes down as you view more images.
- What is the cost of a data breach in terms of manpower?
- As screening goes up, breaches go up.

Discussion of the de-identification process:

- Did you have a formal QC process that involved you verifying the quality of the de-identification process after it was done?
- John Perry: developed a process and a test to make sure it worked, but didn't look at all of the images to confirm it was done without breaches.
- Monitor logs to make sure nothing slips through without automation applied to it. Grab a random 1% and look through the headers.
- Need a more medical model that understands the variability in what we're trying to do
- Partial vs. complete success-field or header
- Catch-22 that you can't crowd-source because there could be PHI
- Build synthetic datasets that have real street addresses in real places that don't match the actual data
- Train a model and release that but not the dataset
- Would need a statistician
- Judy: We are encountering issues that the black box models do not understand. Running experiments on adversarial networks. Surprising findings.
- Amalgamate clinical and imaging data.
- Models have already learned sufficient information to learn age, sex, and race. We don't understand how this happens and maybe they could pick up other identification data.
- We are not trying to hide age, sex, and race. We're trying to prevent the re-identification of a person.
- Increasing the uniqueness of the image data is a threat for re-identification. But if you don't have a database of everyone's fingerprints, for example, it's useless.
- At some point we have to be clear of what we are trying to reidentify and what the practical limits are.
- Clearview.ai

Tasking

- Justin Kirby: report back on what TCIA encounters that is part of their human review processes
- David Clunie: organize report topics in an outline
- Judy: Write up some content (not the overview) on defacing
- TJ and Ying: Can help with defacing

November 9, 2021 Meeting

WebEx recording of the 11/09/2021 meeting

Slides from 11-09-21 meeting

Discussion of threat models:

- David recommends this as a good resource on threat models: https://www.routledge.com/Guide-to-the-De-Identification-of-Personal-Health-Information/El-Emam/p/book/9781466579064
- Another example of a threat model is at https://arxiv.org/pdf/2103.08562.pdf.
- There is a large body of literature on threat models.
- Characterize what kind of attacks you want to consider.
- The ability to demonstrate that there is a flaw in the system. Reasons: Embarrassing the data custodian, demonstrating a new attack method, notoriety.
- A key characteristic of an attack is that it is publicized.
- It is not always prudent to assume that the attacker's resources are limited--free cloud credits. Price of being totally public.
- We don't need to question whether there will be an attack on a public dataset. There will be.
- Risk of reidentification based on the data.
- What are the risk thresholds? Need a reasonable standard. The HIPAA privacy rule says "no reasonable basis to reidentify the individual." Are we looking at maximum or average risk? Think about it from the perspective of the easiest to identify record.
- It's reasonable to assume a .05 probability.
- Prosecutor risk: identify a known individual. We know the individual is in that dataset. Journalist risk: adversary does not or cannot know the target is in the dataset.
- Do you assume everything is publicly available? No. You need to decide what you want to protect–the patient.
- Risk from an insider is high.
- Question: how many people are leaving information that could be used to reidentify? Answer: Manufacturer and model–leaving it demonstrates that different equipment is used.
- Do we see leaving this data in as increasing our risk? No, leaving it in reduces bias. Correlation between quasi-identifiers may not be obvious. Batch effect.
- What is the true utility of the data vs. the risk?
- Risk models–what to include and what not to include. If we limit it to what we need right now, like scanner model and vendor, we may make the data less useful for research later.
- What is the statistical use of the data you've retained?
- We're not normalizing our processes of removing and curating data. Would normalization mitigate risk?
- Is there a reasonable expectation that no one will figure out the geographic location of a scan? The task of the attacker isn't necessarily to determine the geographic location but reidentification to the individual level.

Action for all:

- Think about what we should say about this topic in the report.

December 7, 2021 Meeting

WebEx recording of the 12/07/2021 meeting

Dr. Fred Prior's slides

David Clunie's slides

Discussion of de-facing with Dr. Fred Prior:

The need for de-facing and the risks of faces in images.

Presentation by **Fred Prior**, the leader for investigating this issue for a long time.

- A person's photograph is PII and PHI.
- This question started in 2009 during the caBIG initiative.
- Can average humans match a photograph to an MRI facial reconstruction?
- It was determined in the neuroimaging field that this was a problem.
- Today you can get free MR data. You can get free tools to do the 3D reconstructions. You can use readily available facial recognition data to face data on social media.
- The conclusion has been drawn in the scientific community that you can recognize a photograph through digital data.
- However, the conclusion has not been so drawn by law. HIPAA does not require that images be de-faced.
- OCR/Civil Rights that oversees HIPAA has a rule that if the covered entity does not have evidence that the recipient of the de-ID'd data has the ability to reidentify it, it does not have to be de-faced. It's still a concern by legal people. It's just a matter of time. It will probably change in the next version of the HIPAA regulations.
- TCIA has come to the conclusion that within TCIA, we will compile a complete list of the collections that contain faces. All brain cancers, most head/neck, and radiologic examinations of the head.
- De-facing can hide the data of scientific interest.
- We are focused on the brain tumors we can fix.
- We're not worried about the collections that contain occasional faces. We'll need to develop a search algorithm to find all of these and clean them up. Reject this part of the study/series if it's new, delete or de-face if existing. We are keeping them as restricted access until they can be de-faced. Future research to give us a method to do this.
- We upgraded our data use agreement that says you will not re-identify. State the reason why you want the data. We are not judging these reasons, just recording them.
- We created Masker, our own de-facing software. We have prototyped the pieces and used them on two collections but they're not ready for production. Need to integrate this with POSDA. The curation process work in progress. Sometimes Masker doesn't find the face and curators have to manually add the bounding box. Masker is reversible, so we are open to criticism.
- This came about because of a project from this past summer to explore the de-facing algorithms that are available. We implemented them and they didn't work so we created our own.

- We focused on FSL_deface and MRI_reface from Mayo. It replaces the real face with a generic face.
- All of the existing tools used NIFTI, which is a problem.
- Technical issues impede the full implementation of this. Edge cases where we can't find the face.
- Adding de-facing to the TCIA process takes a lot of time.
- Collections that must be restricted access have been communicated to NCI.
- The goal is to have a communicate to our community tomorrow about doing the switchover with the data.
- Longer term the complete TCIA database will be de-faced.

David's slides and discussion:

- Insider attack
- Outsider attack
- Schwartz et al experiment is difficult to extrapolate from but has a lot of impact on the common understanding of the capabilities of AI.
- Real-world numbers: How many people in the US with gliomas to compare with? 100,000 over a 5-year period, 65 median age.
- If we train on reconstructions, how can you quantify reconstructions?
- Literature needed to inform the HIPAA regulation writers.
- Neuroimaging bias in this context. The wrong conclusions can be reached quickly.
- HIPAA has the statistical arm and the 18 elements arm. Peoples faces may not be useful in a specific context that can be shown statistically.
- Do we need to write a paper or do an experiment? Can we do experiments with data that could risk its status?
- We need a statistical expert who is familiar with quantifying reidentification risk.
- Judy would love to run experiments.
- Could we do experiments with TCIA data? License, data use agreements...
- Judy said she was able to get her IRB to approve experiments with head-neck data.
- Create a sub-group of this task group to plan these experiments.
- Can we apply Facebook's re-id algorithm to a "very large" site (with enough patients to achieve statistical validity on its own)? Federated experiment that aggregates findings to avoid risking re-identification of any individual institution's data. Could get approval for something like this.
- Brian Bialecki: coordinating centers that know the location of the sites submitting the data, even if they don't retain the patients identities, could be used to reject/narrow matches, since a match to a different geographic location than the catchment region of the site could be assumed to be a false positive.