Nueuroimagery datasets are locked away, the HIPAA/HITECH privacy rule pressures researchers to not share datasets or use cycles to ensure the correct paperwork trail. For many research collaborations original source data is not needed, instead sanitized or HIPAA limited datasets can be used instead.

However the tools and techniques used to sanitize datasets are limited in the current NIRTC repository, and do not integrate into the workflow tools employed by various research labs.

This proposal as a first step in comprehensive redaction for large-scale collaboration seeks to extend the Extensible Neuroimaging Archive Toolkit (XNAT) project with a workflow for DICOM redaction.

Our redaction process differs from the traditional anonymization process in that original source data is removed and replaced with tokens. These tokens are one-way identifiers, which map back to source data. This allows investigators to share the datasets freely as long as no HIPAA data is present and lets collaborators begin to process the neuroimagery data and determine requirements for protected information. If protected information is required these tokens map back to a XNAT project and a redaction map database, where an investigator can ensure data sharing agreements are in place and release additional rights to view the removed information.

To accomplish this new workflow, we will design and implement;

* Stand alone redaction tool for DICOM images
* Privacy map database
* XNAT workflow to generate and track redacted/limited datasets

This toolset seeks to fit into the current XNAT workflows and typical research lab processes, but increases privacy of data by making sanitization easier and more secure by sharing using a lease privilege model. In addition this redaction process is build on the tenants of legal redaction and evidence thus protecting investigators in compliance matters.

The long-term aim of this research push is to have smarter form of sanitization, instead of the brute force anonymization tools now in place, the end goal is to allow more intelligent and streamlined collaboration. We envision a suite of tools not only to allow redaction at the DICOM level, but also extend our previous work in low redaction to redact entire hard drives, where PHI can be hidden in various sources from file system metadata to data indices. Incorporating low-level redaction and integrating redaction of logical constructs, such as skull stripping, allow comprehensive redaction of medical images which neuroscience desperately needs to share data freely.