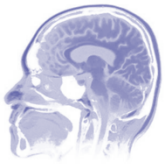
Comprehensive Redaction for Neurological Imaging

Alex Barclay – Computer Scientist[1] Ph.D. Student[2]
Nakeisha Schimke – Ph.D. Student[2]
Dr. John Hale, Ph.D. – Professor of Computer Science[2]

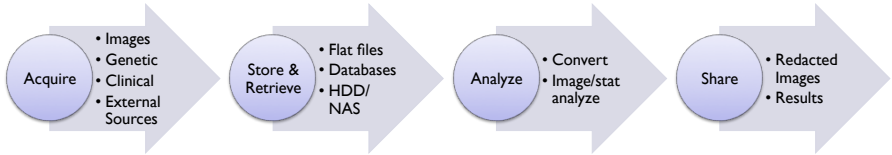
1 – Laureate Institute for Brain Research, Tulsa OK, USA
2 – University of Tulsa, Tulsa OK, USA
Institute of Bioinformatics and Computational Biology



The Problem

- Large-scale studies have huge amounts of data (1PB/3Yrs)
- Data shared must be HIPAA compliant
- Inter-organizational collaboration must be easy
- Data exists as multiple abstractions, and simply removing it from a single layer is insufficient
- Method is needed to specifically redact all layers of PHI

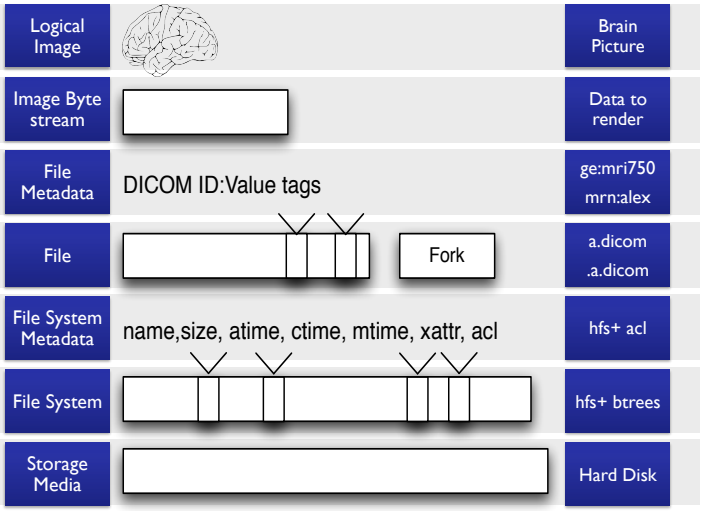
Workflow



Why redact?

Data Stack

- Logical (not architectural or physical) break down of different storage and display components
- Bottom up approach to understand what contains PHI

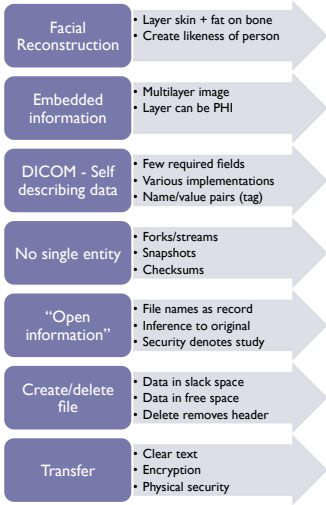


Example

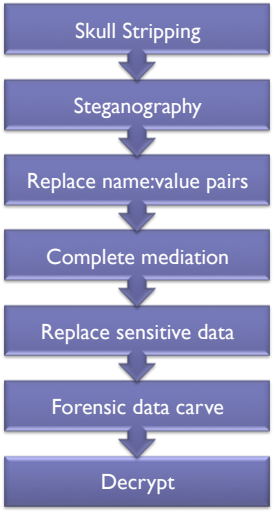
- Mac OS X
- Functional MRI
- Osirix/AFNI viewer

PHI Issues

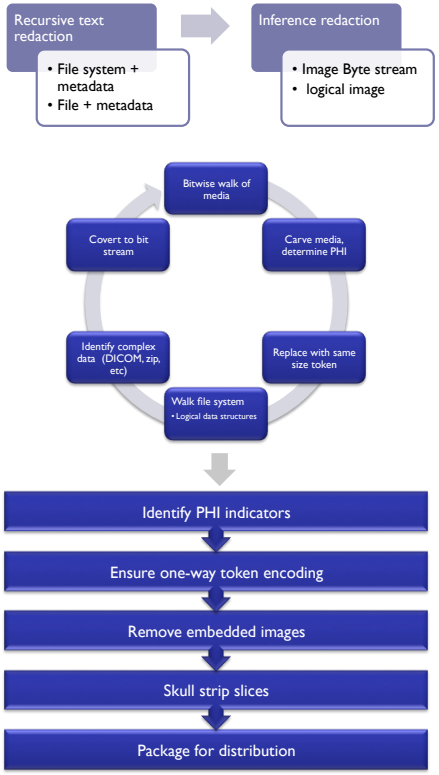
- Partial list of problems by layer mapping
- Complex, need knowledge of neuro study, computer science, and file formats



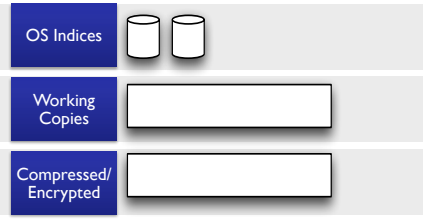
Remediation & Redaction Techniques



Our Redaction Process



External stores



- Data is not just stored in expected file location,
- Depending on application+ file system architecture, data is transferred to working copies
- Operating systems also index data for fast searching
- External data stores when producing images sets require redaction upon the entire media.
- Low level (bitwise) redaction of media is the only way to find all sources of PHI

Discussion + Future

- Comprehensive redaction is combination of recursive redaction at the block and file layer, with additional techniques to find and reduce inferred data
- Based upon body of work and code for legal production
- First step in exploring imaging specific redaction issues
- Expand and bundle current tools, produce integrated tool

References:

Arkfeld, M. R. (2005), *Electronic Discovery and Evidence, Law Partner Publishing, L.L.C.*
Bischoff-Grethe, A., et al. (2007), *A technique for the Deidentification of Structural Brain MR Images. Human Brain Mapping, 28:892-903.*
Carrier, B. (2005), *File System Forensic Analysis, Addison-Wesley Professional.*
G. Manes, L. Watson, E. Downing, A. Barclay, D. Greer, J. Hale, *A Framework for Redacting Digital Information from Electronic Devices,* Proceedings of the 8th Annual IEEE SMC Information Assurance Workshop, West Point, New York, June 20-22, 2007.

Acknowledgments:
William K. Warren Medical Research Institute
TU Bioinformatics reading group
David Greer + Patrick Bellgowan
Photobucket – mri picture