

## **Trusted Image Dissemination (TID)**

Professor Gio Wiederhold (PI)  
Department of Computer Science  
Stanford University

In TID (Trusted Image Dissemination), we propose to provide image filtering capabilities to complement other means of checking the contents of documents. An example is to restrict information contained in images that are part of an electronic medical record to avoid violations of security or privacy.

An increasing amount of information being transmitted over the Internet is in image form. This trend includes medical images used in diagnosis or research. Such information has not been processed in the past with concern for security or privacy and most privacy and security control is oriented towards text. The TID approach will provide an innovative capability based on experience with image database and protecting the privacy of information in databases.

We will extend the facilities we have developed in current security-oriented projects at Stanford (TIHI, SAW) to provide more thorough filtering of medical information, including images containing text. The TIHI effort, supported by NSF's HPPC challenge program, has now built a prototype of a software tool--called a security mediator--that enables legitimate external customers to obtain remote electronic access to medical information residing in a medical institution, while inhibiting the release of contents that should not be released, even when the requesters appear to be authorized. The successor project, SAW, focuses on protecting shared manufacturing data. Image filtering is becoming relevant in manufacturing domains, since much computerized information in manufacturing and business involves images, but security of image contents is not supported within the scope of most research efforts.

Nearly all approaches to security focus on controlling access. Unfortunately, controlling only access requires a perfect organization of the internal data in an enterprise. In many practical cases this requirement cannot be fulfilled, since it implies a radical reorganization of all internal information services. Aligning all internal data to deal with external access privileges is not only costly from the systems point-of-view, but also for all internal users of information systems, who now must file all data according to external requirements that are normally none of their concern. For instance, in a hospital, if some X-rays are to be released for research purposes, then certain identifying marks, used internally to prevent misdiagnoses, must be omitted for external release. In manufacturing, drawings containing proprietary data must be edited if the decision is made to have the parts produced by an external subcontractor.

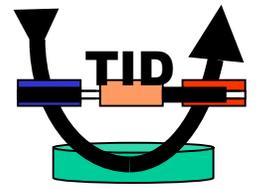
Filtering of images in addition to text is becoming essential, since modern computing has greatly facilitated the use of information in image form. We will focus TID initially on electronic medical information, but the principles are broader.

In related research, we have developed novel means of recognizing features in images, specifically in linking perceptual factors to parameters in wavelet-based analysis of images. We have experimented with a number of tools, mostly based on parameterized wavelets, that can recognize crucial information, such as text in images, and submit it to the content checking rules our base system provides. Initial results are very promising.

The image filtering in TID will rely primarily on wavelet technology. Work has been completed at Stanford that demonstrates the capability of indexing and retrieving images by wavelet transform analysis. The wavelet approach to filtering has been demonstrated to be fast and highly reliable. Its formal basis provides better scalability over more ad-hoc approaches, and has also been easy to transfer among programming languages. TID will especially consider text that is placed within images and hence not accessible to routine processing. We anticipate that electronic medical information transmission over the Internet will be a major area of application, but many other web pages today place text within images to increase the attractiveness of the presentation.

Our objective in TID is to extend our security and privacy protection work to multimedia databases, specifically adding a capability to handle medical images. Our efforts will focus on:

- developing further an existing wavelet-based algorithm for searching medical image databases (content-based image retrieval) and developing techniques to retrieve digital images and relevant textual information from multimedia medical databases.
- Extracting textual information from retrieved images.



- defining rules relevant to the protection of privacy of medical images and implementing them into a security mediator.
- developing techniques to edit medical images automatically in accordance to rules.
- adapting and developing tools for manual editing of digital medical images, specifically to omit identifying information.
- developing further a WWW customer interface for the proposed security mediator.