

# The National Software Reference Library

Barbara Guttman



# Why We Started Collecting Software

NSRL was asked by the Law Enforcement Community (in 1999) to collect and process software, resulting in a data set of known, traceable applications that can alleviate the effort in determining important evidence on computers.

# How We Collect Software

NSRL acquires software through commercial channels, purchasing shrink-wrapped releases or downloading from established distributors and via donations from vetted sources.

Includes: PC, Apple, Linux (and other) software, mobile apps

Currently around 15,000 packages

# Beyond Forensic Use

NSRL is expanding into software identification to support computer security and other applications.

A long-term goal is to provide a standard infrastructure for identifying software, files and other software-based artifacts to support a broad range of technical and social applications.

Our interest is primarily in developing the technical infrastructure for identification.

# Software Identification & Assurance

Block hashing

Approximate matching

Software diskprints: record all changes (registry, memory, network traffic) during installation, use and removal

SWIDtags

National Vulnerability Database

# Stanford Collaboration

NSRL is leveraging existing infrastructure and processes to ingest the Stephen M. Cabrinety Collection in the History of Microcomputing at Stanford University.

Media images, materials photographs and metadata are shared between both organizations.

# Future Plans

More mobile

Automated collection of large software distributions (e.g., Steam)

# Even Farther into the Future

Partner with “real” libraries for collection,  
preservation, and access

# Contacts

Barbara Guttman  
Doug White

[douglas.white@nist.gov](mailto:douglas.white@nist.gov)

[barbara.guttman@nist.gov](mailto:barbara.guttman@nist.gov)

NSRL website

[www.nsrl.nist.gov](http://www.nsrl.nist.gov)