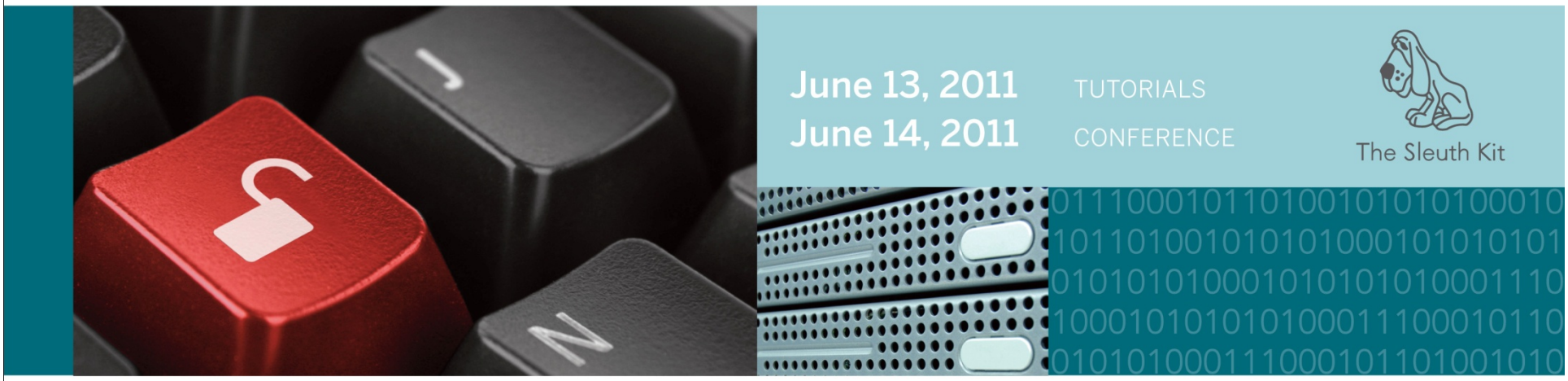
2ND ANNUAL CONFERENCE

The Sleuth Kit and
Open Source Digital Forensics Conference

BASIS TECHNOLOGY



Sleuth Kit and Autopsy 3.0 Update

Brian Carrier
Basis Technology Corp

nference

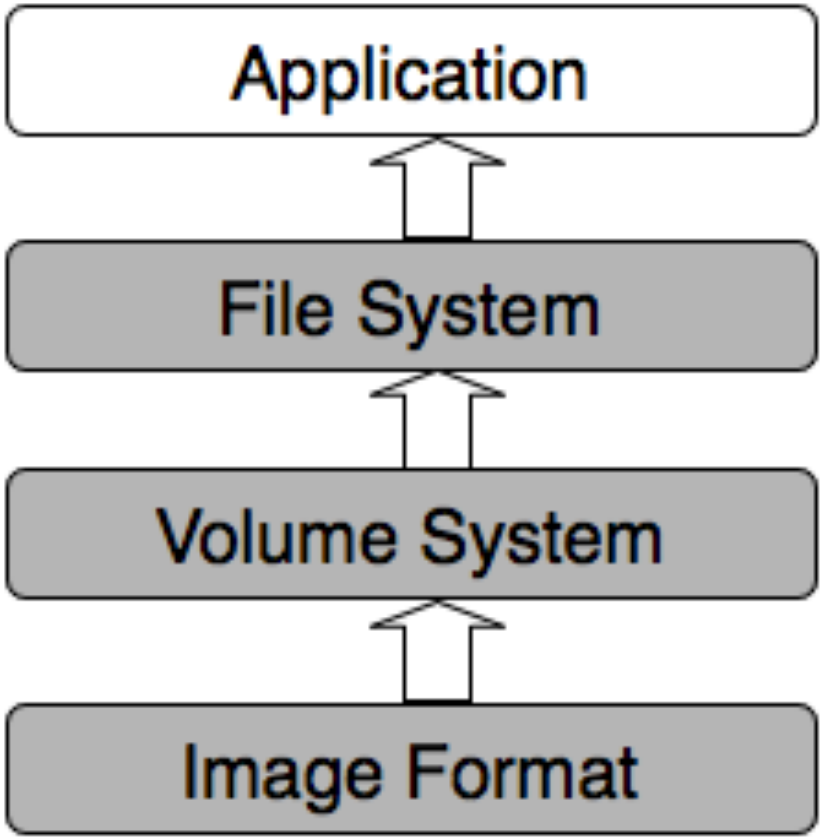
Agenda

- What is TSK
- What's new since last year
- What's planned for this year
- Autopsy 3.0
- Hadoop Prototype Framework

What Is The Sleuth Kit?

BASIS TECHNOLOGY

Open source software that allows you to forensically analyze disk images and local drives.



Scenario

- You have a disk image and want to look for specific files.

1. TSK will auto-detect the image format
2. TSK will auto-detect the volume system and layout:
   - What sectors are allocated to partitions
   - What sectors are not allocated to any partitions

Scenario (contd.)

3. TSK will auto-detect the file system type and can search for your file (even if it is deleted)
   - Analyzes the directory hierarchy in file system.
   - Identifies files that have been marked for deletion.
   - Searches for “orphan files” that no longer have a name.

Command Line Tools

- Original method for using TSK
- Currently, over 25 different tools
- Mmls example:

```
# mmls tsk1.img
     Slot    Start        End          Length       Description
00:  -----   0000000      0000000      0000001      Primary Table
01:  -----   0000001      0000062      0000062      Unallocated
02:  00:00   0000063      0032129      0032067      NTFS (0x07)
03:  00:01   0032130      0064259      0032130      DOS FAT16 (0x06)
```

Fls example

- Lists the files in a directory.

```
# fls -o 63 tsk1.img
r/r 4-128-4:     $AttrDef
[…]
r/r 3-128-3:     $Volume
d/d 29-144-6:    dir1
d/d 31-144-1:    dir2
d/d 34-144-1:    RECYCLER
v/v 19920-144-1: $OrphanFiles
```

Library

- All of the command line functionality, in a C/C++ library.
- More efficient to use when processing a full disk image.
- Reduced overhead:
  - Load general file system data only once
- Full API docs and sample programs exist.

Library Quick Start (New School)

- Create a C++ class that extends TskAuto.
- Implement the processFile() method
  - It will get called for every file in an image.
- That's it!

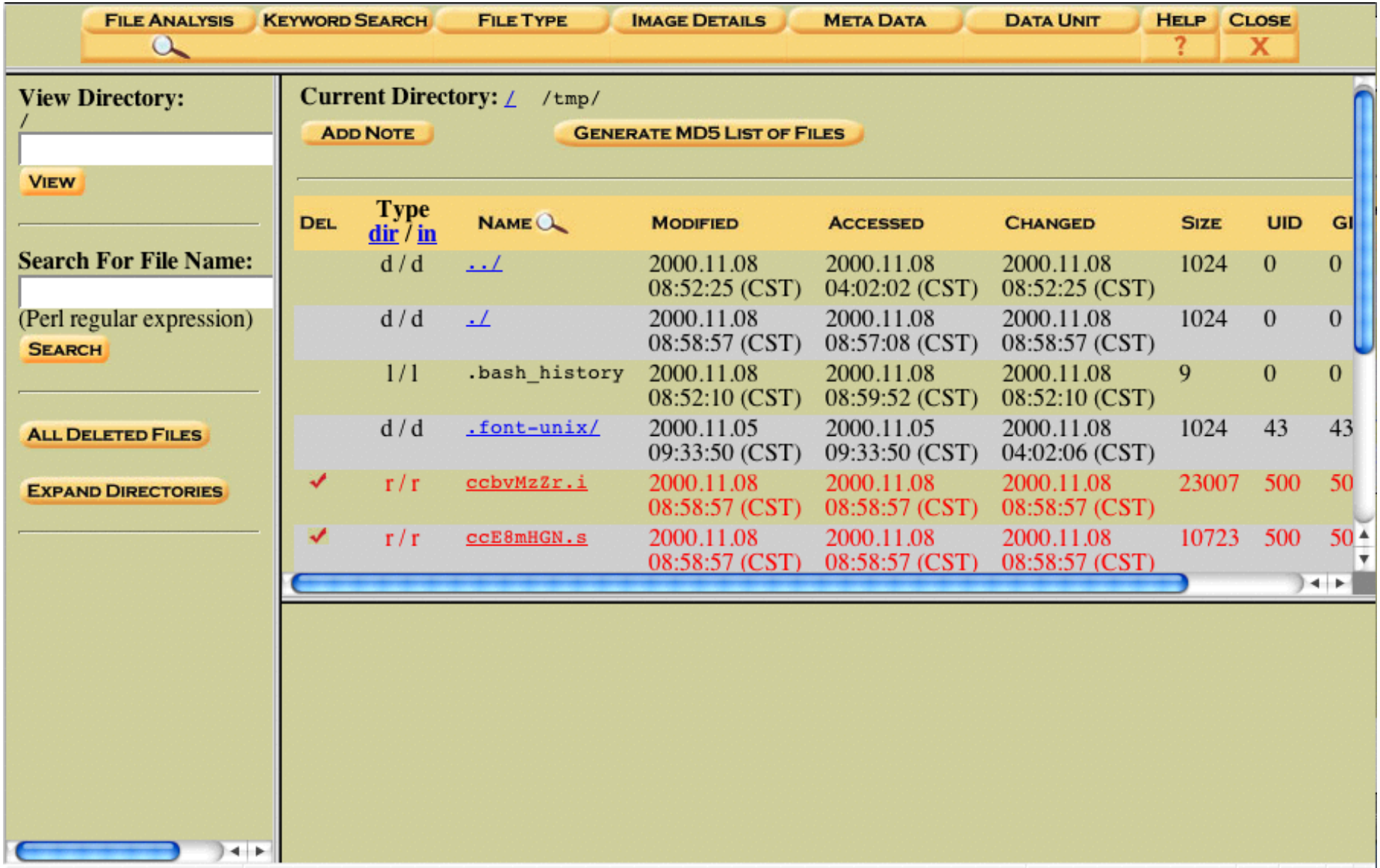
SQLite Database

- Use ‘tsk_loaddb’ or library to dump file system data to SQLite database.
- Open database in your program using the language of choice.
- Reduces the number of required cross-language bindings.

Autopsy

- Original graphical interface to TSK
- First released in 2001
- HTML-based interface:
  - Runs TSK command line tools
  - Parses output and adds HTML tags
- Does not use the library interface.

Autopsy 2



Lots of other tools…

- Open source tools
- Commercial tools
- Bootable CDs

- Refer to wiki.sleuthkit.org for full listing.

What's New Since Last Year?

TSK Changes

- Releases 3.2.0 to 3.2.2
- New TskAuto class
- SQLite database output
- RAW CD Format
- Performance
- Better data corruption handling
- New tools & functionality

SQLite Database Overview

BASIS TECHNOLOGY

- Tables store file system metadata:
  - Image_info: Image size and type
  - Vs_info: Describes each volume system
  - Vs_parts: A row for every volume
  - Fs_info: Describes each file system
  - Fs_files: A row for every file
  - Fs_blocks: Map files to their blocks
- Does not store any file content.

New 3.2 Tools

- Tsk_recover:
  - Extracts files from disk image.
  - Creates directory hierarchy in local file system.
- Tsk_comparedir:
  - Compares local directory hierarchy to disk image.
  - Useful for detecting rootkits and testing.
- Tsk_gettimes:
  - Equivalent of 'fls –m' on all file systems.

What Has Yet to Be Released

Multi-threaded Support

- Threads allow systems to take advantage of multiple cores at the same time.
- Locks were added to TSK.
- Works on all platforms.
- None of the released tools use multiple-threads.
- Code is in the public source code repository and will be included in 3.3.0.

C++ Wrappers

- New C++ classes wrap C functions and structs.
- Same functionality, but more data encapsulation.
- Helps to enforce thread safety.
- Code is in the public source code repository and will be included in 3.3.0.
- Sample programs and documentation exist.

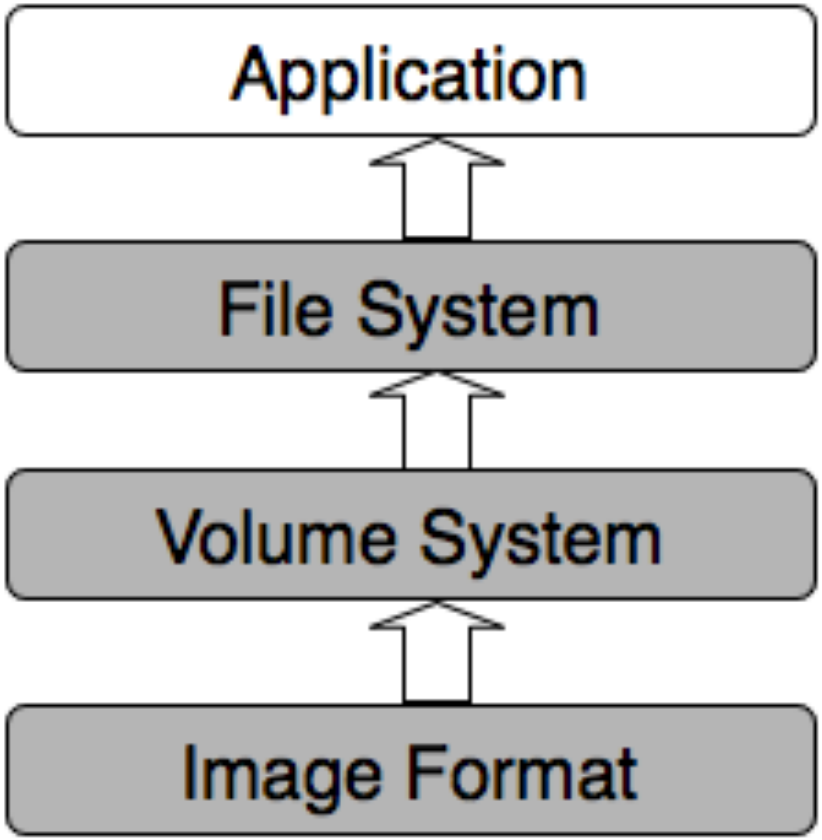
JNI Java Binding

BASIS TECHNOLOGY

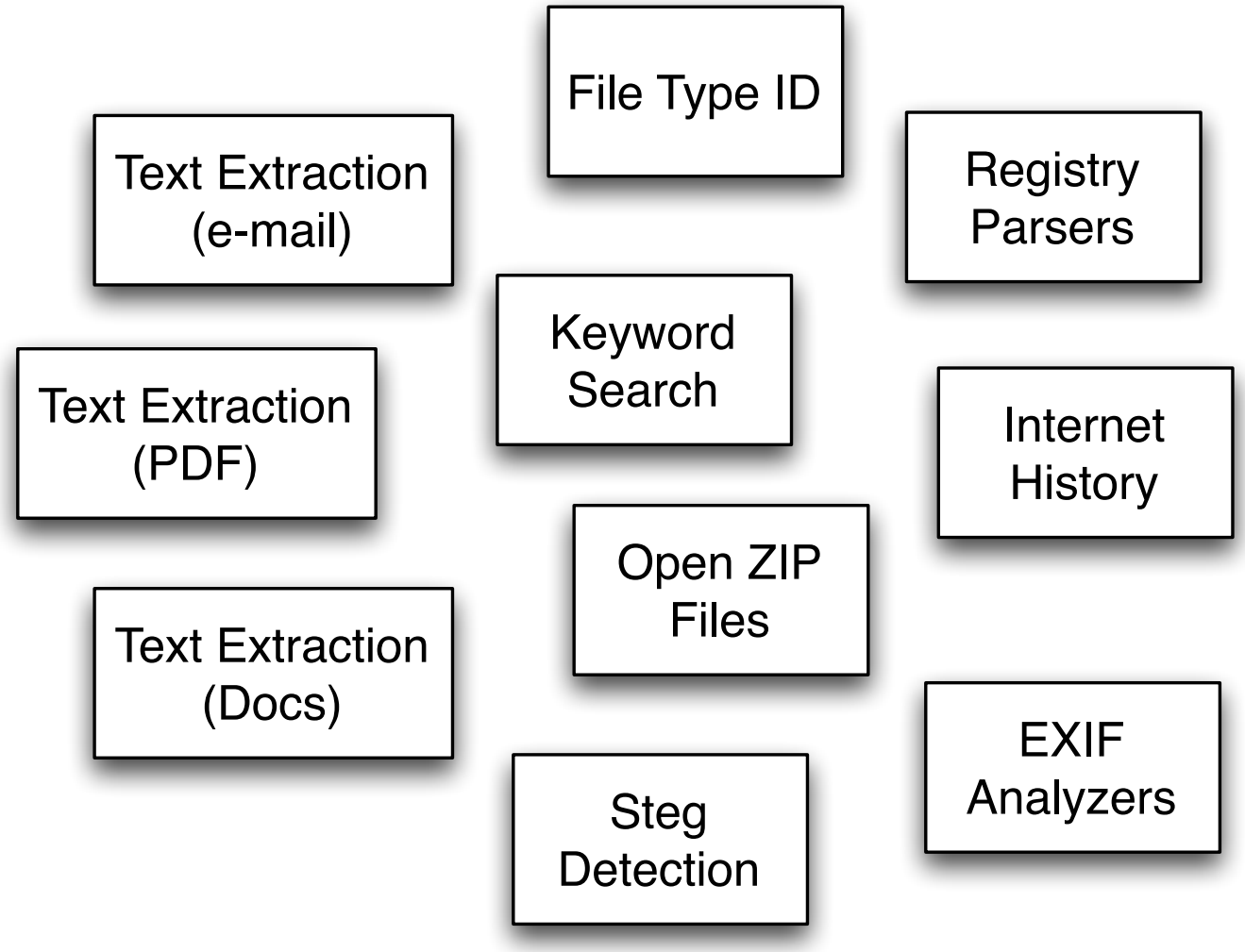
- Allows Java programs to use TSK C library.
- Can create SQLite database with metadata.
- Can call library functions to obtain file content (not stored in database).
- Code will be checked into public repository.

What is Planned to be Released

Application-Level Framework



Application-level Examples



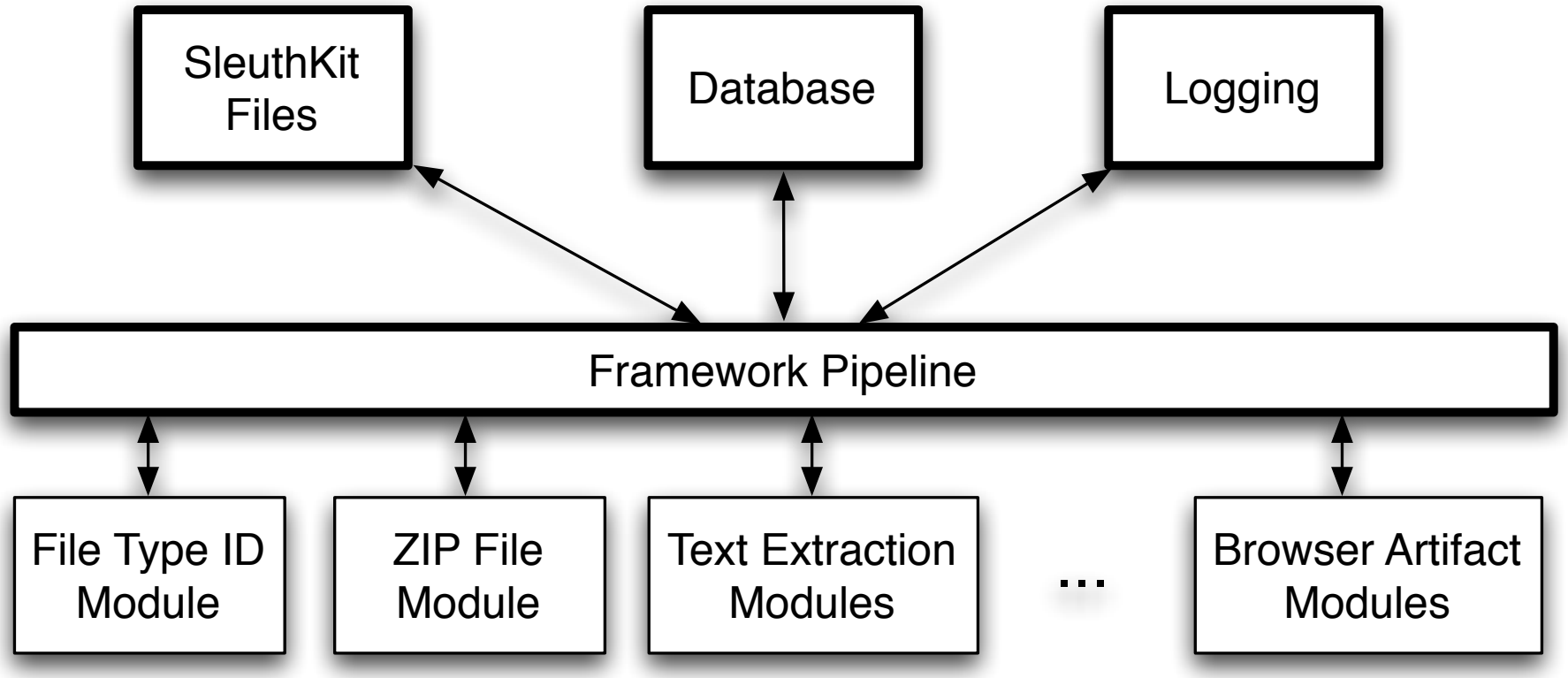
Framework Basics

- Pipeline:
  - A series of plug-in modules
  - A file is analyzed by running it in the pipeline
  - Defined with an XML file
- Database:
  - Stores analysis results
  - Can also be used to store file metadata
  - SQLite or a client-server database

Modules

- Dynamic Library Plug-in Modules:
  - Has access to file content and metadata
  - Has access to results from previous modules
  - Can write analysis results to blackboard
  - API: analyze(File)
- Reporting Modules
  - Run after all of the files have been analyzed
  - Creates output report
  - API: report()

Framework



Help Will be Needed

- If you build it, they will come.

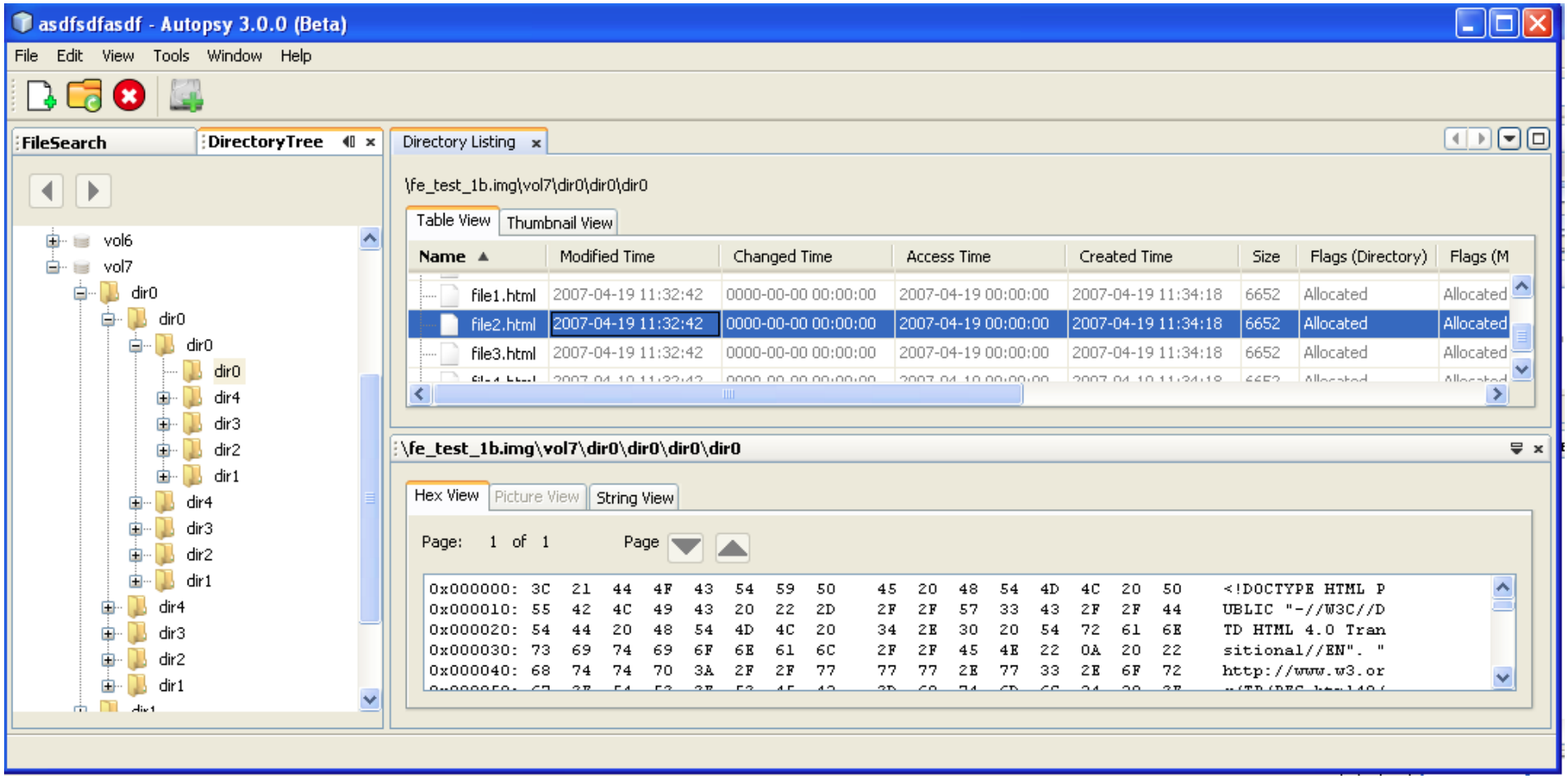
- We can't create all needed modules.
- Ask other tools to write TSK modules:
  - Internet artifacts
  - Registry
  - …
- We'll provide docs for doing this.

Autopsy Version 3

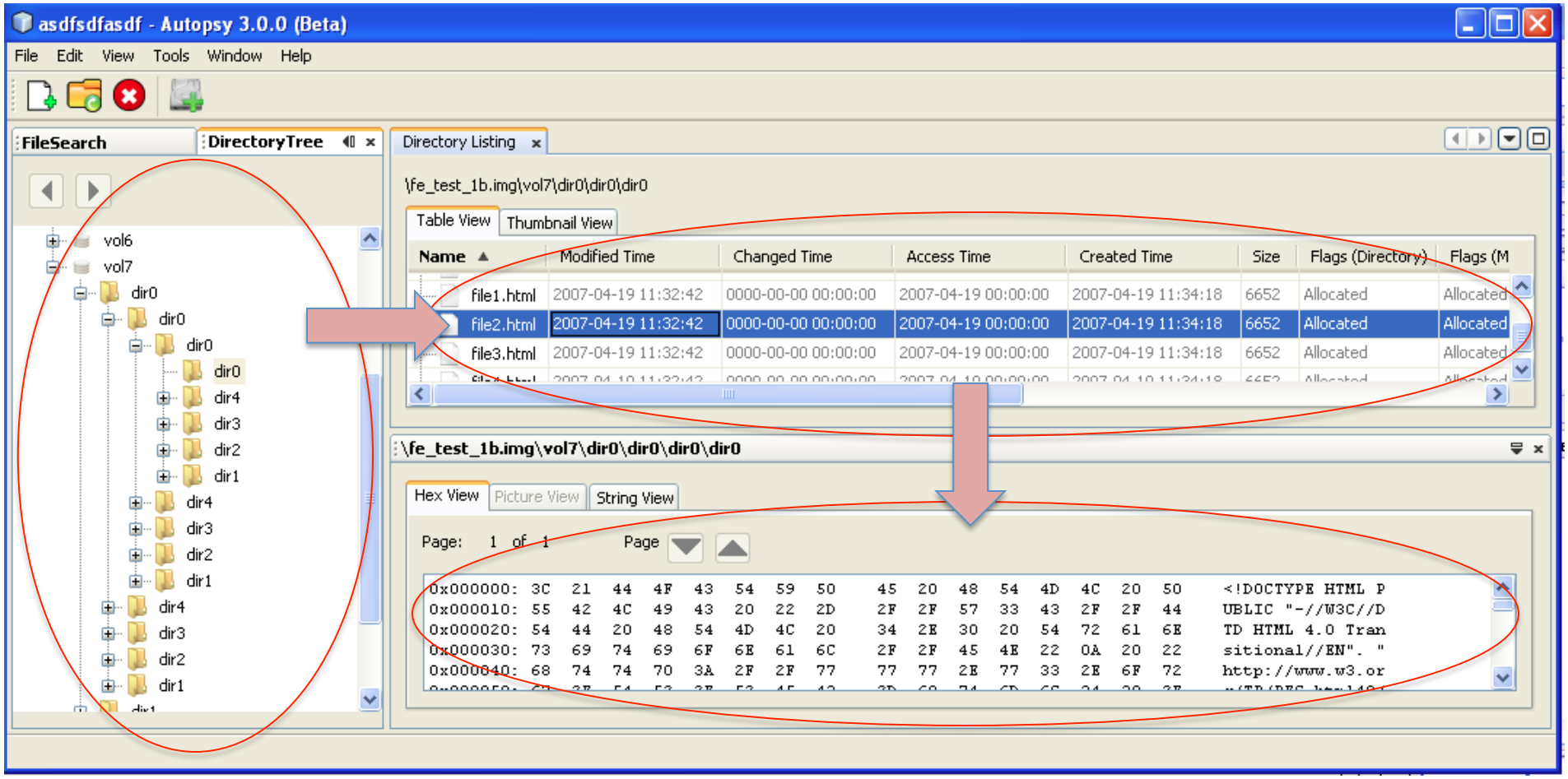
Autopsy 3.0 Basics

- Java GUI will run on multiple platforms.
  - Currently, only Windows
- Based on Netbeans Rich Client Platform.
  - Allows for easy module integration
- Will allow us to leverage Lucene and other Java open source software.
- Uses SQLite database and JNI bindings.

Autopsy 3 Screen Shot



Modular Design



Current Features

- Left-side:
  - Directory Tree
  - File search (by name, times, size)
- Upper-right:
  - Table listing
  - Thumbnails
- Lower-right:
  - Image viewer
  - Strings view
  - Hex dump

Plug-in Analysis Module 101

- Left-side can be used for interface.
- Access disk image and file data using internal Autopsy services.
- Save results as “Netbeans Nodes”.
- Push nodes to upper right area.

Planned Features

- Keyword search
- Timeline analysis (log2timeline)
- Hash database integration
- Bookmarks
- …

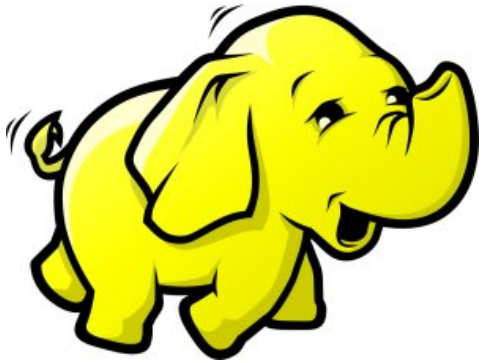
- First beta release will be in July.

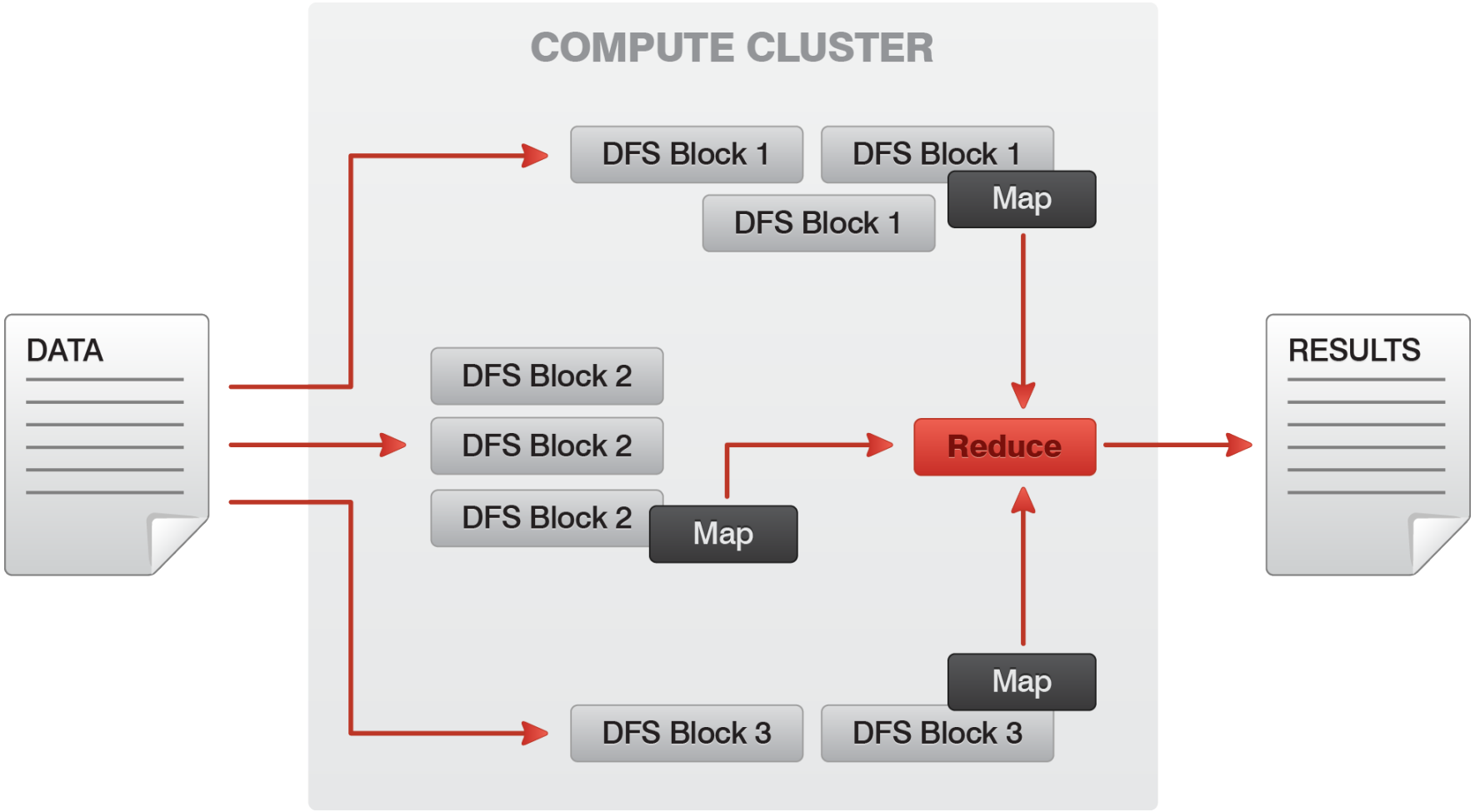
Hadoop

Basics of Hadoop

BASIS TECHNOLOGY

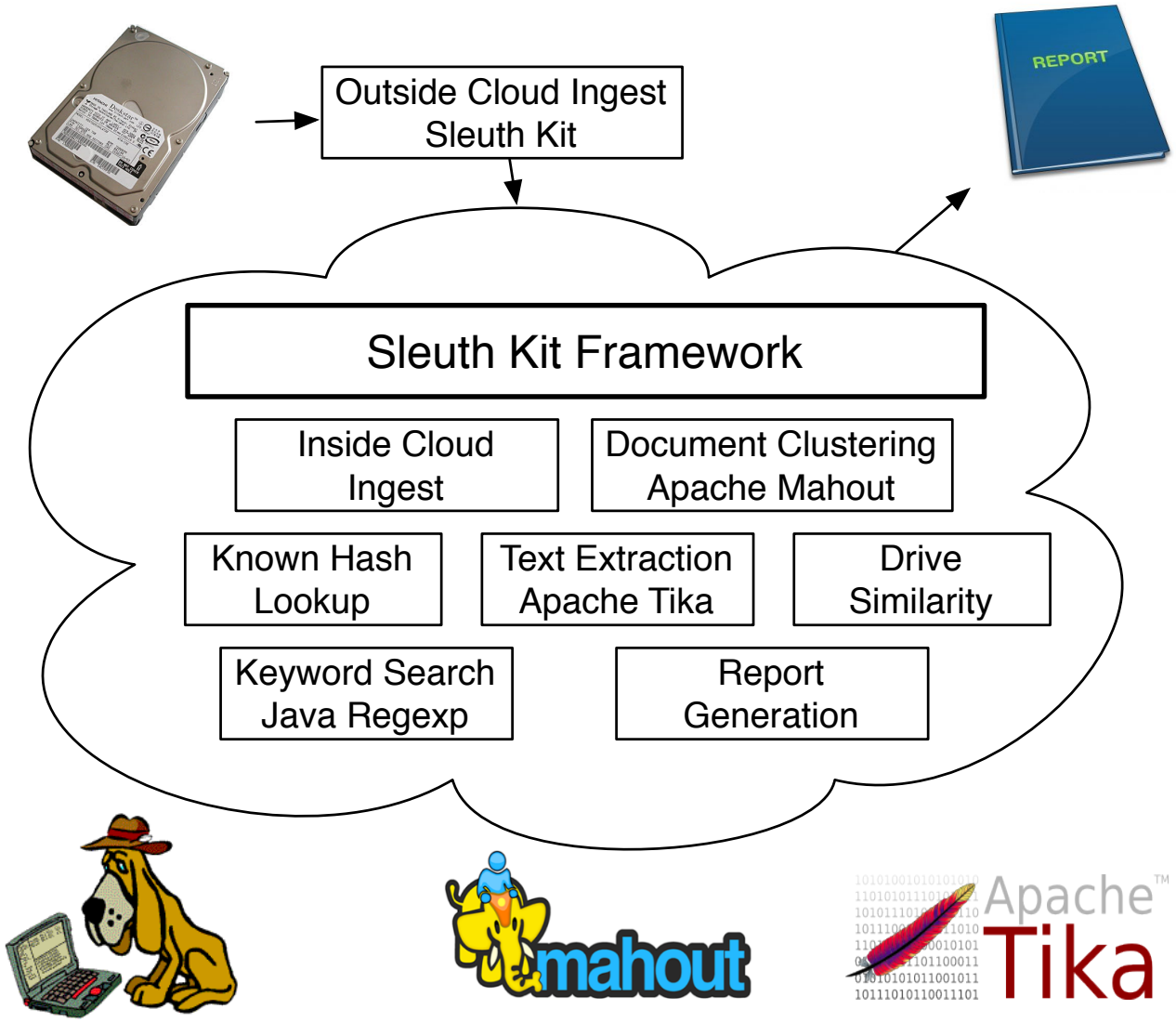
- Open source Apache project for distributed computing.
- Based on papers that Google has published
- Provides (among other things):
  - Scheduling among thousands of nodes
  - Distributed and localized storage
  - Resilience if nodes fail
  - …
- To get these features, you must formulate your work as a series of “MapReduce” tasks





Prototype Framework Project

- Joint project with:
  - 42Six Solutions
  - Lightbox Technologies
- Funded by US Army Intelligence Center of Excellence (USAICoE)



Next Steps

- Still working on prototype.
- Still collecting numbers on performance.
- Will be released as open source later this summer.

Thank You!

For more information

Visit www.basistech.com

Write to info2011@basistech.com

Call 617-386-2090 or 800-697-2062