elaman
GERMAN SECURITY SOLUTIONS

SCOOTY
Speech Classification Online and
Offline Technology

SCOOTY – Speech Classification Online and Offline Technology

SCOOTY Overview

What can SCOOTY be used for?

SCOOTY is a speech analysis system. In a process from signal detection, recording, processing and filtering to evaluation, SCOOTY is part of the evaluation by analysing audio files for the following features:

• speech/non-speech detection
• language identification

The core parts of SCOOTY executing this analysis are called "classifiers". Based on a client / server architecture, SCOOTY provides a software environment for:

• training
• configuration
• classification
• production

How is SCOOTY used?

The process is the following: After classifier training and configuration, SCOOTY produces classification results without further user intervention. Training and Test result data as well as configuration data are stored in the data resource (a kind of an archive) and are recalled for production.

Production results are also saved in the archive and concurrently shown on screen to allow monitoring of progress and results.

Who uses SCOOTY?

Everybody working on speech and language is a possible user of SCOOTY software. Typical users of the software are people working on huge amounts of speech data coming in. This can be a university working on the structure of speech and languages, such as dialects, accents, and language history. A benefit from the use of SCOOTY can also be drawn by research institutes in medicine treating with diseases in the production and perception of speech.

What is the benefit when using SCOOTY?

Using SCOOTY, large amounts of speech signals can automatically be processed and sorted according to their characteristics.

The performance of the system depends on the training data that are fed to the system. Speech files can be sorted according to the characteristics, giving an indication on the spoken speech. Eventually, an indication can be given also on the dialect, depending on the training data.

The accuracy for the classification of a language depends on the involved languages and especially on the training material that is used for the estimation of the classifier. Out of 1000 files for classification, usually 600 files can be classified correctly, according to the spoken language.

Does the structure of SCOOTY fit into my environment?

The SCOOTY software is based on a client/server architecture consisting of the following parts:

• SCOOTY client(s) – graphical user interface
• processing server – signal parameterisation and classification
• archive server – data management
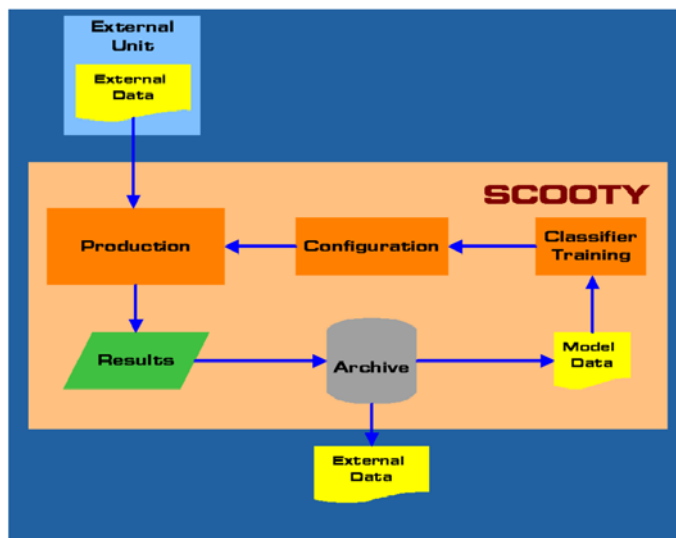
Each of these software parts is available as single user version or network version, and each of them is available for different operating systems. They can be combined freely and platformindependently via TCP/IP, so they are highly adaptable to system and application requirements.
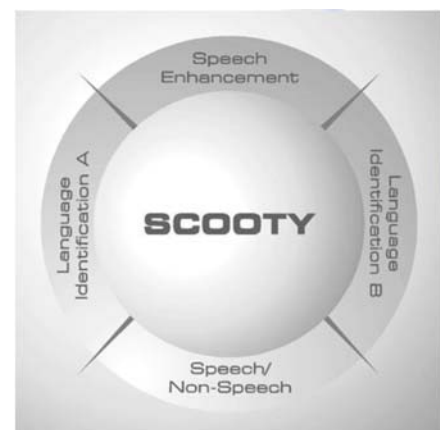
Classifiers

How does SCOOTY use the classifiers?

SCOOTY performs the classification with the following classifiers:
• a speech/non-speech classifier
• two language classifiers using different algorithms estimating the probability for the involved languages.

Furthermore, SCOOTY **enhances** the **speech** quality of the input signal to improve the reliability of the classification results.


SCOOTY applies cascaded classifiers and the signal enhancement to the input signal containing WAVE files. The classifiers are equipped with an extensive parameter set. The parameters are adjusted during classifier training to gain optimum quality of classification results.
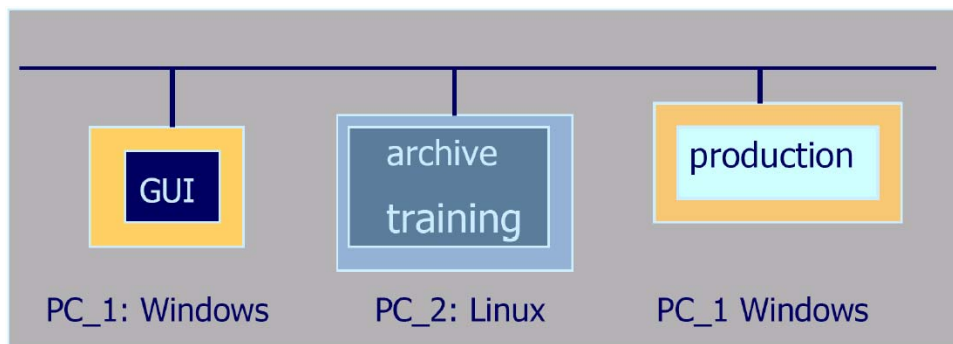
System Configuration Examples

Example 1:

SCOOTY runs on 1 PC or notebook and is used by 1 user.
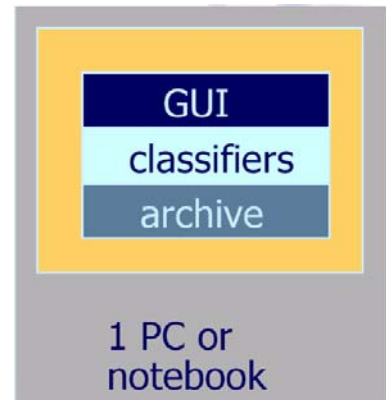
**Example 2:**

SCOOTY is distributed over 3 workstations connected via a network. Still it is used by 1 user but the workstations perform different tasks, such as classifier training and production.

Example 3:

SCOOTY is used by multiple users working on different tasks at different places. Operators at place A work on classifier training and designing different SCOOTY configurations, while other users work directly on the production results. A network connection between places A and B is not necessary.
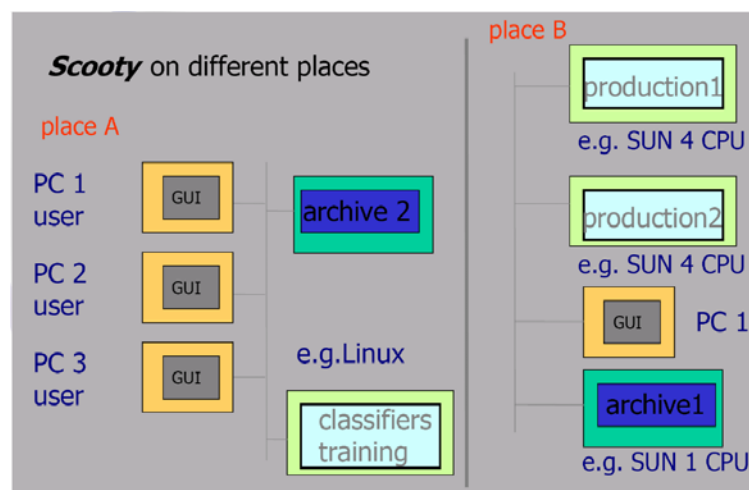
Classifier Parameterization and Training

How can SCOOTY find "my" languages?

The SCOOTY classifiers are trainable to allow them to be adapted to individual customer requirements. The results of the training process are special model data, which are later used during classification. Best classification results are achieved with customer training and test data.

SCOOTY is equipped with a complete training and test environment for the classifiers. This covers a transcription tool to generate label files (they contain information about whether an audio file contains speech, and if so, about the language) and a graphical user interface for easy classifier training and testing.

Classification Process Configuration

Is SCOOTY flexible to meet my requirements?

SCOOTY allows the use of the available classifiers in different combinations:
• cascading of classifiers and signal enhancement
• exclusion of unnecessary classifiers
• use of differently parameterized classifiers
Different combinations and cascades of classifiers can be summarized to signal flows. The following figure shows 2 valid cascades, each specially adapted to a certain scenario:
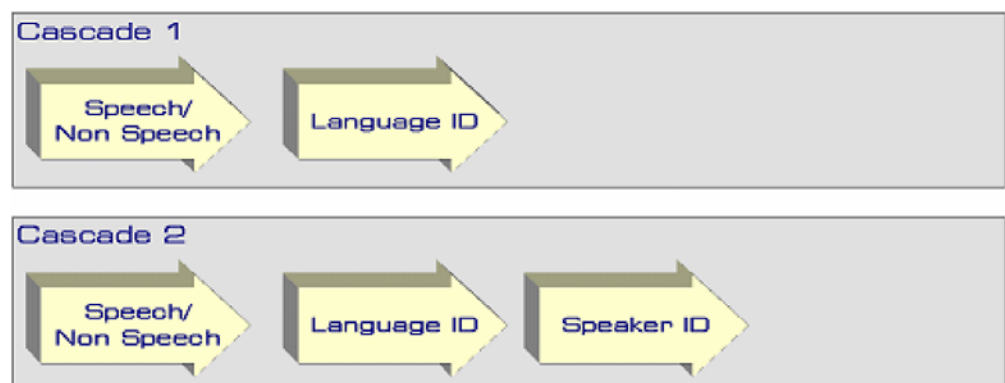

Production

How does SCOOTY work?

After starting the production, SCOOTY reads its input directory and processes the input files in a hierarchical order. Each input file represents a task for SCOOTY. SCOOTY processes them in the order of priorities. For each task, SCOOTY produces result files and additionally displays the results on the screen. No further user intervention is required, but the user can observe the production progress and status.

During production, SCOOTY selects one of the configured signal flows (cascades) for each task. The selection depends on information provided by an external control unit. The file then runs through a cascade of classifiers assigned to the selected signal flow. In the cascade, the file is finally analyzed. This concept allows SCOOTY to analyze signal files from different sources.

The figure shows the data flow during the production:
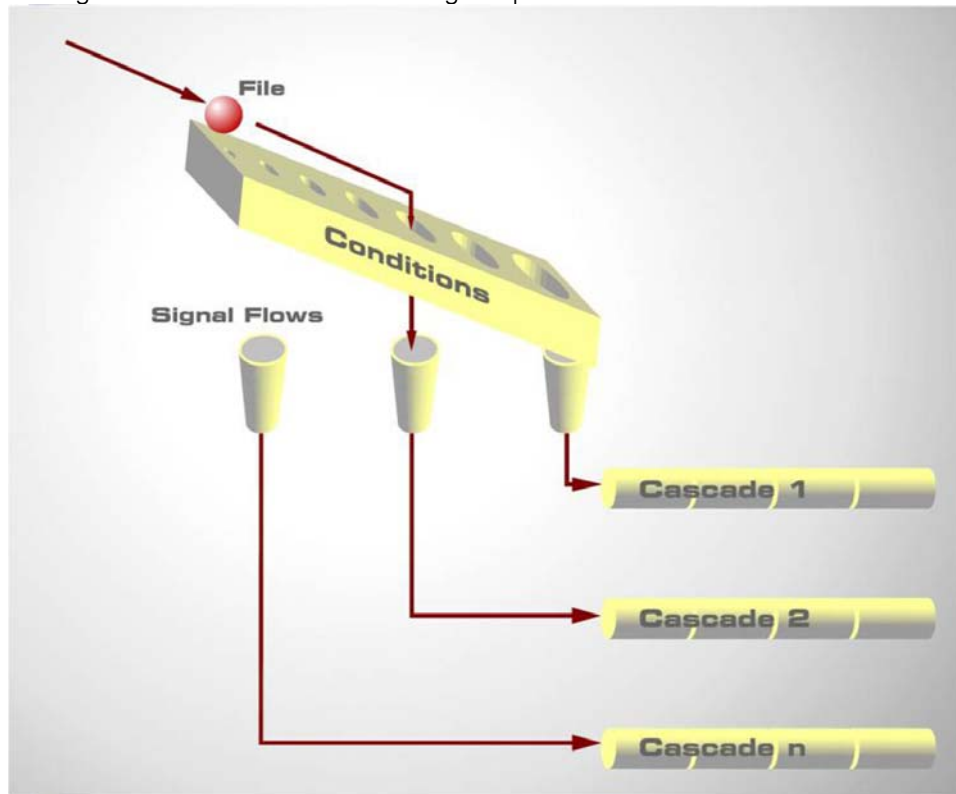
A processing server coordinates the classification process. It encapsulates classification functions in a way that they are externally available as services.

Result Retrieval and Display

**How can I view and further process SCOOTY's results?**

Classification results are:
• Detected speech
• Identified languages

SCOOTY stores the results, the protocol of all activities and processed files in an archive:
• A work log of each classification (text format)
• Label and score files of each classification (linked information in text and XML format)
• Training and test results, model data, test results (text and XML format)
• Signal files (WAVE format)

The LogViewer is an extensive retrieval tool to recall the results from the archive. Other result files are additionally stored in an output directory from which they can be taken by other processes. SCOOTY provides all results for extensive statistical computing. The archive allows the user to search for all types of data.

Additionally, during production temporary results are displayed on screen:
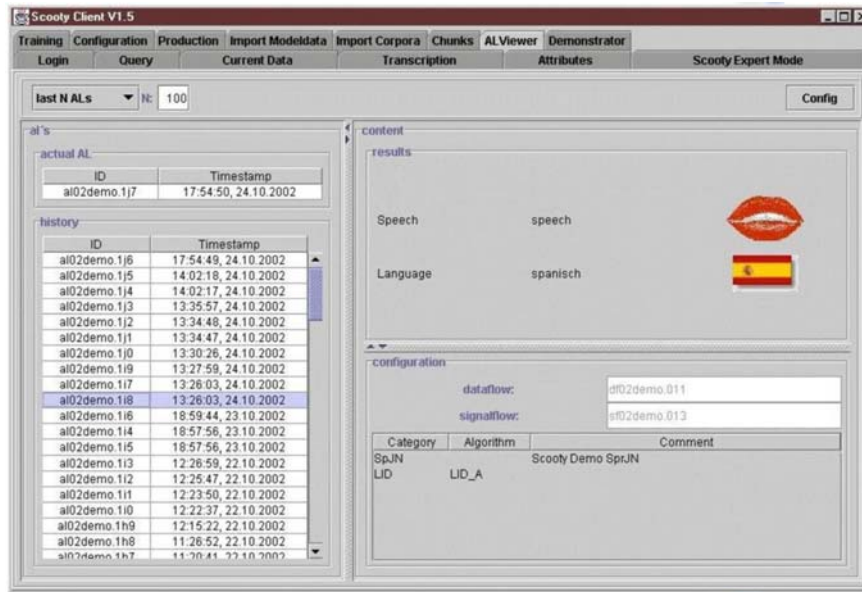
Data Management

Isn't it difficult for me to handle all these complex data?

All information is provided with attributes, after which it can easily be searched and sorted. SCOOTY provides a user-friendly graphical user interface for comfortable and fast data retrieval and viewing.

A central data handling resource (the archive) is the pool for any type of information. All relevant data appearing in the SCOOTY workflow are stored on the archive server. The SCOOTY archive assures traceability of processes, data consistency and cross-reference consistency. It stores:
• training, test results and model data
• configuration data
• production results
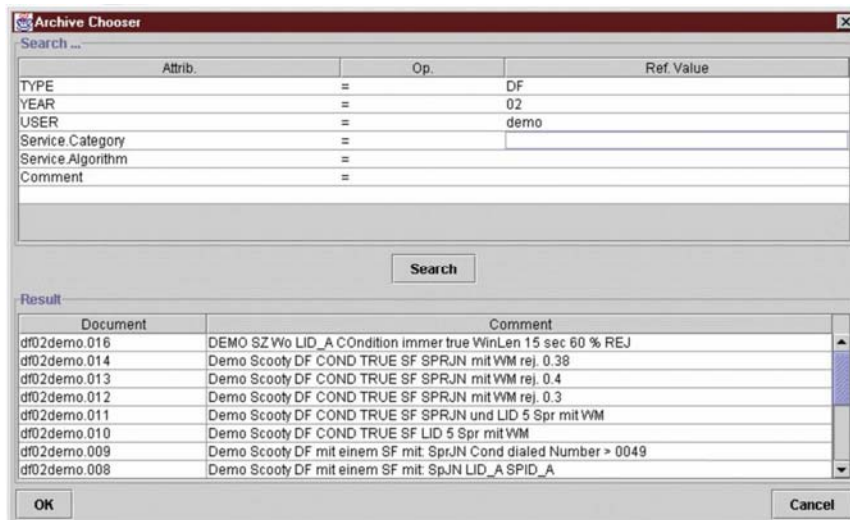
Further Developments

I need different information from the audio signal. Will there be more classifiers?

Currently, the following classifiers are under development:
• word spotter
• topic spotter based on sound files
• text visualisation and classification
• speaker identification

Furthermore, there will be the opportunity of text analysis of audio files.

Feature Summary

The following list summarizes the features of SCOOTY:

• System for automatic signal classification
• Server client architecture
• Use of standard hardware
• Scalable
• Extensible with new classifiers
• Trainable with customer data
• All algorithms for training available
• Improvement of the quality of classification
• Full support preparing data for training
• Powerful archive for offline job analysis
• Support of statistical computation based on classified data

Do not hesitate to contact us if you need more detailed information.

Further Features and Services

Scalability	SCOOTY can be run on systems ranging from a single user – single computer (all components installed on one computer) up to client server networks (components distributed over several computers).
Extensibility	The software is based on a modular concept. All modules are available separately and for different operating systems. For details see the SCOOTY pricing schedule.
User Training	User training is available for training and configuration of classifiers, setting up production configuration, and evaluation of results according to customer specific requirements. The training level depends on customer requirements. The training can be held at customer's site or ours, using example data or data from customer environment. Contact ELAMAN for details.
Consulting Service	ELAMAN helps and consults on classifier training and testing, or, if required, performs the classifier training and test for customers.

Software Support   The software licenses include free software support inclusive of all updates for 1 year.

Customer Documentation The system is equipped with a user manual an installation instructions. The software provides context sensitive online help. All documentation is available in English and German.

System Requirements

Platforms (all program components)

• Windows 2000, Windows XP
• Linux (S.u.S.E. 7.x and higher)
• SUN Solaris 8.x

Archive Server:

• Disk space: > 100 Gbyte, depends on usage
• RAM: > 256 MByte
• CPU: one processor machine

Server for Classifiers

• Disk space: ca. 40 GByte for training
• RAM: > 512 MByte
• CPU: two or more processor machine

Client (Graphical User Interface)

• Normal PC or notebook
• Disk space: 30 GByte
• RAM: 256 MByte
• CPU: ca. 600 MHz Pentium

Interfaces

• Input: file based (Text-/WAV files), header file with information must be specified
• Output: file based (Text, XML)

Performance Data

Hardware dependent, e. g.: Pentium PC 1-2 GHz
• a single LID classifier faster than realtime

Company Principles and Policy

Technology
… in development and company management is state-of-the-art,
and represents only the best.

Quality
… in all areas of our company is regarded as the almost requirement
for risk-free and successful cooperation with our customers,
and business partners.

Market Position
… we are the specialists in the field of signal and data processing
as well as pattern recognition, and we are glad to face competition.

Colleagues
… form the roots of our company, and give the performance required
for maintaining and building the technical base, and close
personal cooperation we have with our clientele.

Growth
… we strive toward a healthy, stable foundation at home and
abroad.

Services
… are comprehensive and complete. As a full-system company
we offer standard equipment, systems, and services.

Trust
… in the relationships to our business partners, and within our
own company forms the basis of our business.

If you would like further Information about ELAMAN,
or would like to discuss a specific requirement or project, please contact us at:

Elaman GmbH
German Security Solutions
Seitzstr. 23
80538 Munich
Germany

Tel: +49-89-24 20 91 80
Fax: +49-89-24 20 91 81
info@elaman.de
www.elaman.de