Podmining® - SAILing the multimedia web

Mark P. Pfeiffer
VP Business Development and Communications
SAIL LABS Technology AG is a leading global provider of software-based speech and language technology for multi-media analysis solutions.

Facts & Figures

- Founded 1999 in Vienna, Austria, Europe
- 99.8% Austrian owned
- Research and Development entirely in-house
- Customers: Worldwide coverage
- Current headcount: 20 employees + consultants
- Clients in all continents
Podmining

Podcasts, webcasts & internet clips contain important information sources for industries and governments.
Products & Services

Key Products
- Media Mining System
- Text Mining System
- Podmining

Key Features
- Indexing and archiving of data from video, audio and text sources
- Automatic speech recognition (speech-to-text)
- Speaker and gender identification
- Topic and named entity detection, story segmentation
- Format normalization

Key Services
- Consulting service and technical workshops
- Research, development and implementation of customized solutions
- Maintenance, upgrades and support
Podmining challenges

- No magic bullet available off the shelf...
  - Podcasts are VERY diverse in domain and vocabulary
    - Cooking, Politics, DIY, Astrology, Astronomy, Brain Surgery, Tourism, Kinky Toys, Society Chit Chat, Children's Entertainment, Oil, Economy, Collateral Trading, Adult Entertainment, Art of the 17th Century, Martial Arts, Explosives Handling, Music...
  - Audio quality also has a very broad range
    - Professional recording environment to home studio with 5.95$ mic
  - Speakers are mostly not “trained speakers”
    - Anchor speakers are the minority in this business!

- High flexibility needed in terms of Language Model
  - Vocabulary is broad and needs adaptation for special domains

- Budget
  - Eventually they wanted a solution for reasonable price and we had to be able to pay the rent
How it works
Technology Concept by Pluggd
How they manage to repay...
SeeHere

Technology Preview: Search Within Shows Using Pluggd's SeeHere
SeeHere uses an intuitive heatmap interface to let you search inside a video and jump right to the parts that interest you. To get started, just type a concept into the search box on the video player below. (Note: although SeeHere results will appear immediately, you may need to wait for the video to download before you can jump ahead to a hotspot. The progress bar shows how much of the video has downloaded.)

InDigital - A Bit of Everything!
Published on May 18th (271 days)
In this episode, Jessica compares three 411 services that won't cost a cent and then takes a cameraman for the action spots minded for a spin. We check out a Linux based media player and Hahn tries out a gadget for those that need the Internet everywhere they go. And before the Canucks complain... Hahn needs to apologize to all Canadians for not knowing that Ottawa is still in the Stanley Cup hunt.

How Pluggd's SeeHere works:
Less relevant: information, media
More relevant: online, web, sites

InDigital - A Bit of Everything!
May 18th (271 days ago)
In this episode, Jessica compares three 411 services that won't cost a cent...
THANK YOU!

Contact

SAIL LABS Technology AG
Mariannengasse 14 · 1090 Vienna · Austria
Ph +43-1-580 95-0 · Fax +43-1-580 95-580
info4speech@sail-technology.com
government@sail-technology.com
www.sail-technology.com