

PISA-M

Peer-to-peer based, dynamic and scalable
Infrastructure for Search Appliances on
Multimedia information

ISTI-CNR, Pisa, I (*F. Rabitti, R. Perego*)
Max Planck Inst. Informatik, D (*G. Weikum*)
University of Padova, I (*M. Agosti*)

Contact:

fausto.rabitti@isti.cnr.it

Dec. 19^o, 2005

Information Day on "Audio-Visual Search Technologies"

Problem

- Most of existing data on the **Web** is **multimedia** (images, audio/video, music, etc.). More so in the future.
- Future multimedia search will be embedded in applications such as wireless services via mobile phones, MP3 player, etc. (**search appliances**)
- Existing search engines do not **index multimedia content** (only associated text/links)
- The centralized Web search model cannot handle the space and performance requirement for searching multimedia content (need of **scalable and distributed search structures**)
- Multimedia contents change very often (need of **dynamic search structures**).
 - For example videoblogs, photoblogs, and podcast
- Traditional **crawling algorithms**, where objects are "pulled" from their containers, are inadequate for highly dynamic content.

Dec. 19^o, 2005

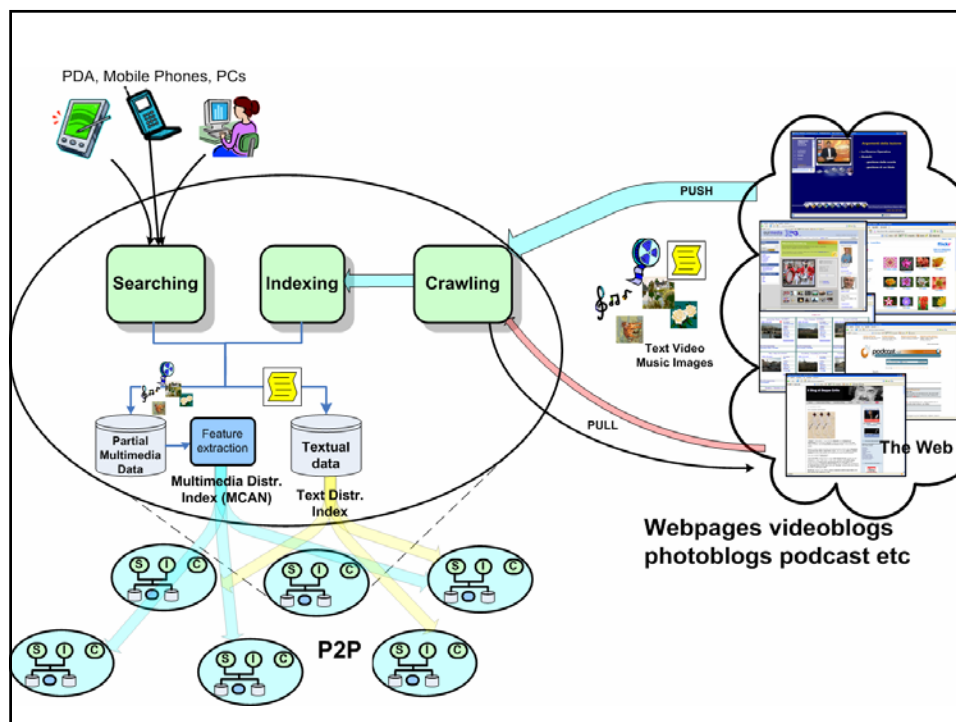
Information Day on "Audio-Visual Search Technologies"

Proposal: New technological infrastructure for MM Search Engines

- **Media specific analysis and understanding** (automatic feature extraction and content classification)
- **Scalable, dynamic and distributed index structures** supporting similarity search (MCAN)
- **Cross media analysis** in query answering: combining evidence (extended by annotations) from different media indexes using the similarity paradigm
- **Support of distributed push-based crawling** (where containers are asked to publish and “push” information to the search engine) together with the traditional pull-based crawling.
- **Towards a distributed search engine architecture** (Peer-to-peer based), as opposed to today parallel search engines within a centralized Web data warehouse.
- **Scalable dynamic caching techniques** to enhance performance
- **Search infrastructure suitable also for IPR-protected MM content providers** (news feeds, music, images, etc.), independent from DRM servers.

Dec. 19^o 2005

Information Day on “Audio-Visual Search Technologies”



The project

- Target application:
 - *MM search for mobile application generating summarized interactive objects as result*
 - *Self-organizing search networks of independent nodes*
- IP or Strep? Depending on industrial involvement
- Partners:
 - Research
 - Industry: mobile device and mobile communication/service providers
 - Content providers: (news, music, images, etc.) ???

Dec. 19th 2005

Information Day on "Audio-Visual Search Technologies"