

Position Statement for *Audio-Visual Search Technologies*

Wray Buntine
Helsinki Institute for Information Technology

December 19, 2005

1

Infrastructure for Search

To encourage vibrant business activity around search and related media services: develop the services and standards to support organisations and businesses providing access to community media.

Broad spectrum content analysis: provision of cheap/free infrastructure services for analysis to allow widespread use of semantic-based methods.

Integrated personalisation and reputation infrastructure: support for social networks, citations and web links integrated with content analysis and personal preferences.

Standards: use of and extensions to standards such as OpenSearch RSS, Common Query Language (CQL), Open Archives Initiative (OAI) Protocol for Metadata Harvesting, SiteMaps, etc.

Embedding the search interface: support integration of media access and search, *e.g.* embedding search with the above facilities directly into blog and news readers, and in music or video P2P systems.

2

Broad Spectrum Content Analysis

Packaging and support for services to allow non-specialists to configure, adapt and deploy in arbitrary subject areas.

Tagging: automated tagging for key dimensions of content analysis:

Genre or Style (*e.g.*, service and repair, consumer information, FAQ),
Named Entities (*e.g.*, people, businesses, cities), and
Subject Category (*e.g.*, Tourism, Aerospace).

Resources: packaging and reuse of broad spectrum public linguistic and semantic resources such as WordNet, Wikipedia, DMOZ Directory, etc., to support the above.

Services: web services and access methods to support external specialised media and search systems.

3

Integrated Personalisation and Reputation Infrastructure

Consumer oriented personalised search (specifically, “find what I like”) integrated directly into the media access platform (P2P music system, news reader, etc.).

Personalisation and community: *e.g.*, community access, friends lists, recommendations, and preferences to support personalisation of search.

Reputation systems: domain specific reputation systems combining web link and text data, citations, and results of content analysis.

Link database: easy access to link text and URL data to allow specialised search services to employ topic specific link analysis without a need for their own global crawl.

Data Valet: personalisation service data in a sharable format so that individuals can manage personalization across providers.

4