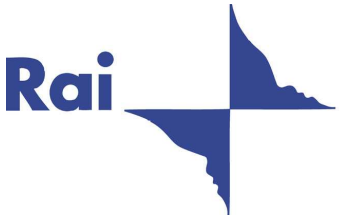


UNIVERSITÀ
DEGLI STUDI
DI TORINO
ALMA UNIVERSITAS
TAURINENSIS

A Generalised Cross-Modal Clustering Method Applied to Multimedia News Semantic Indexing and Retrieval

Alberto Messina, Maurizio Montagnuolo

RAI – Centre for Research and Technological Innovation

Madrid, April 23rd 2009

18th International Conference on World Wide Web – WWW 2009

Contents

- ✓ Background and motivations
- ✓ State of art
- ✓ Architecture
- ✓ Core technologies
- ✓ Experimental evaluation
- ✓ Conclusions

Background and Motivations

- ✓ News aggregation is not a new concept
  - ☞ Yahoo! News
  - ☞ Google News
- ✓ But mono-modality is a serious limit in these applications
  - ☞ Only news published on the web
  - ☞ Aggregation based on textual content only (e.g., spoken content is not considered)
- ✓ Objective of our work: make users able to use multimodal sources of information about news events
- ✓ Overcome limitations of mono-modality with ***multimodality***
  - ☞ Cross-modality (i.e. of things coming from different sources)
  - ☞ Multimediality (i.e., of things coming through different media)

State of art

- ✓ RSS (Really Simple Syndication)
  - ☞ Information overload effect
- ✓ [Huang and Webster 2004] uses keywords to characterise RSS contents
- ✓ [Yan et al. 2007] uses k-means to cluster RSS referenced documents
- ✓ [Banerjee, Ramanathan and Gupta 2007] uses wikipedia to augment the quality of clustering
- ✓ [Xu, Wang et al. 2008] integrates web-casting text and broadcast content for sports event annotation

State of Art

- ✓ Information mashup
  - ☞ RSS aggregators
    - ☹ Mono-media data
  - ☞ Information Mash-Up
    - ☹ Multi-media data but mono-modal distribution channels (e.g. WWW)
  - ☞ Web-casting text and broadcast content integration
    - ☺Multimodality
    - ☹ Users are “passive” observers
    - ☹ Mono-directional, “one-to-many” relations
- ✓ Topic Detection and Tracking (TDT)
  - ☞ Automatic identification, linking and accessing to topically related information items within heterogeneous, real-time news streams
    - ☹ Symmetric similarity among data
    - ☹ Not hierarchical relationships

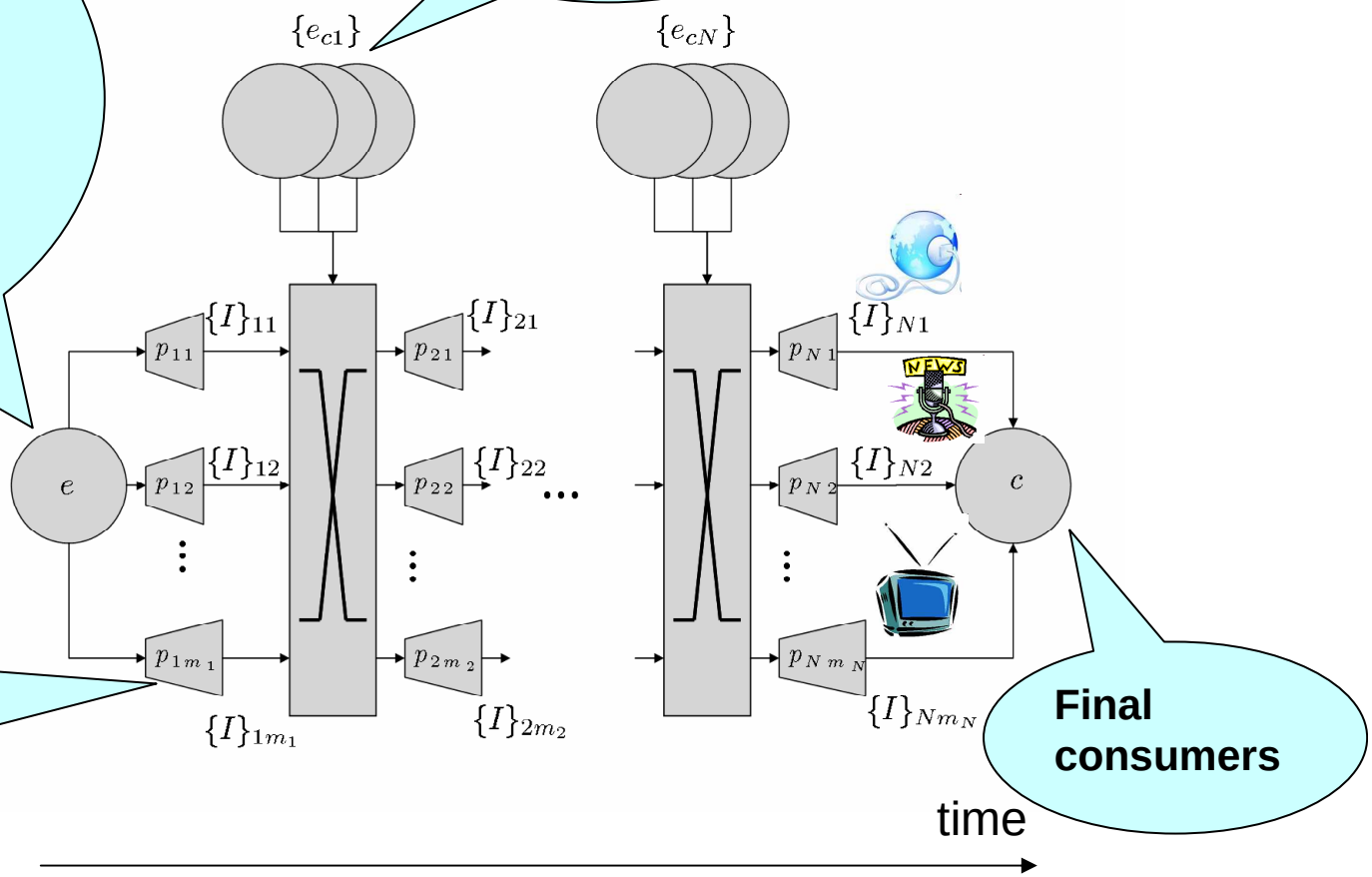
Our proposal

- ✓ Unsupervised framework for acquisition, segmentation and aggregation of broadcast news
- ✓ Cross-modal hybrid clustering between RSS items (published on web) and newscast items (broadcasted)
  - ☞ Asymmetric semantic affinity among items
  - ☞ Hierarchical, many-to-many relationships
- ✓ These two ingredients make our system fully ***multimodal*** (in the defined sense)

Conceptual Model (1)

A news event e is *"any relevant fact happened at some time and place, along with all the information elements generated during its occurrence".*

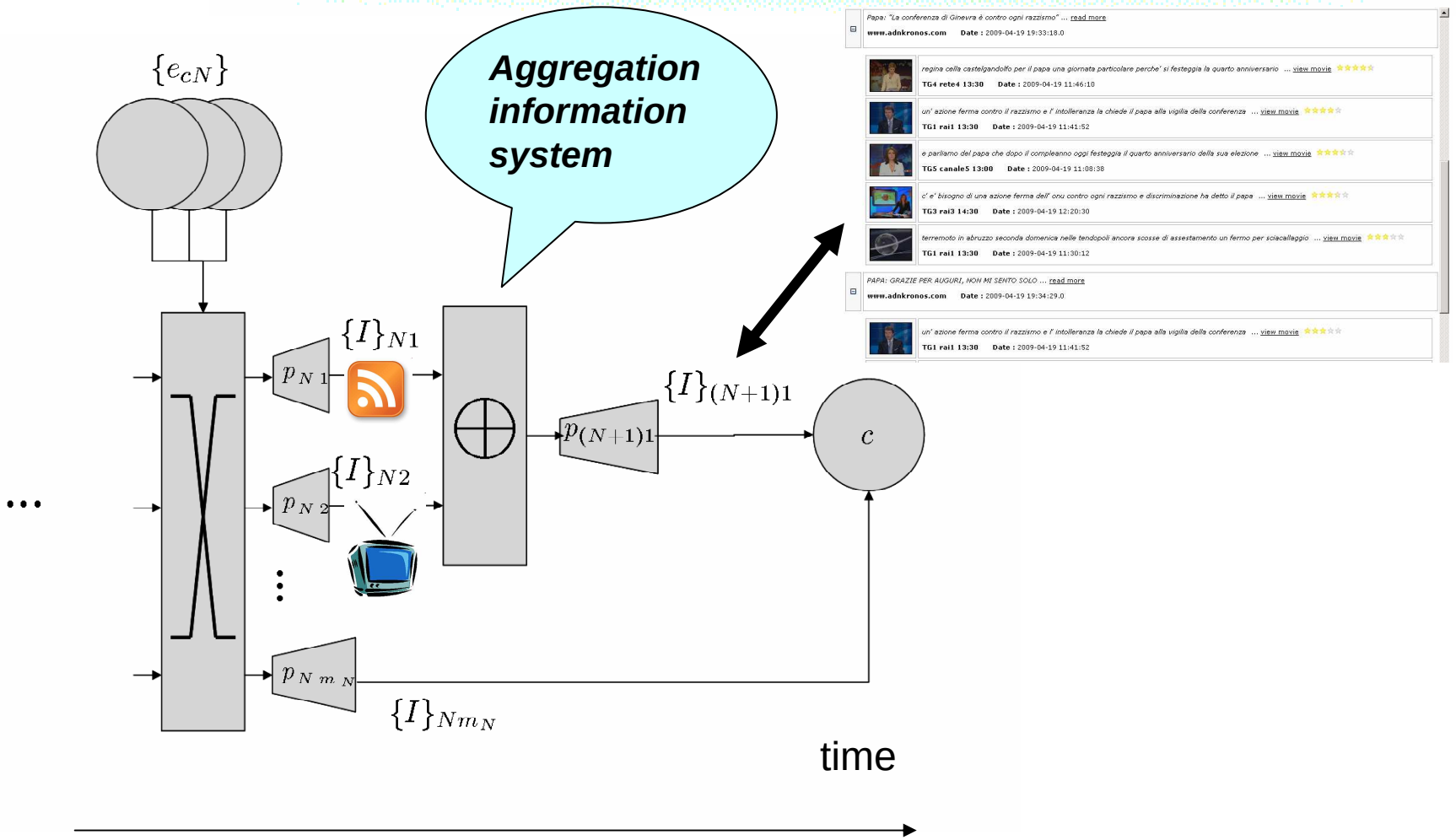
Contextual news events



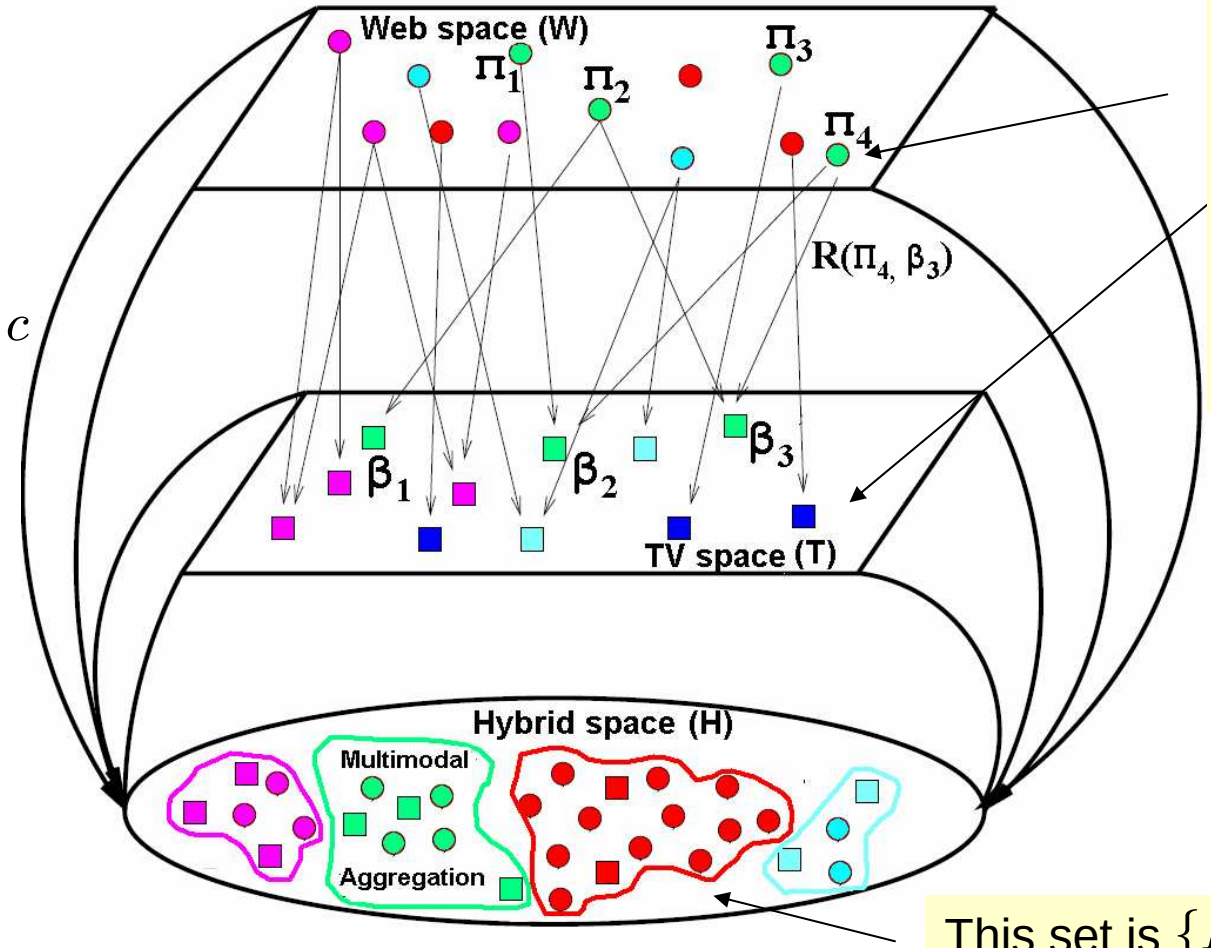
***Information providers* (e.g. press agencies, broadcasters)**

Final consumers

Conceptual Model (2)



Scenario



These sets are $\{I\}_{N1}$ and $\{I\}_{N2}$ sampled in a given time window. We call them Tv and W resp.

This set is $\{I\}_{(N+1)1}$

Core technologies: hybrid clustering

- ✓ Semantic affinity concept
  - ☞ i is semantically aggregable to j if the fruition of i satisfies expectations about j
  - ☞ Degree of semantic affinity $s_{i,j} = R(\pi_i, \beta_j)$
  - ☞ $\forall \pi_i \in W \forall \beta_j \in T_v$ we estimate $s_{i,j}$
  - ☞ We arrange these coefficients in an affinity matrix $\mathbf{A}$ whose rows collect the degrees of semantic affinity of a certain element of W

$$\mathbf{A} = (\mathbf{S_1} \ldots \mathbf{S_m}) \in R^{m,n}$$

$$\mathbf{S_i} = (s_{i,1}, \ldots, s_{i,n})$$

Core technologies: hybrid clustering

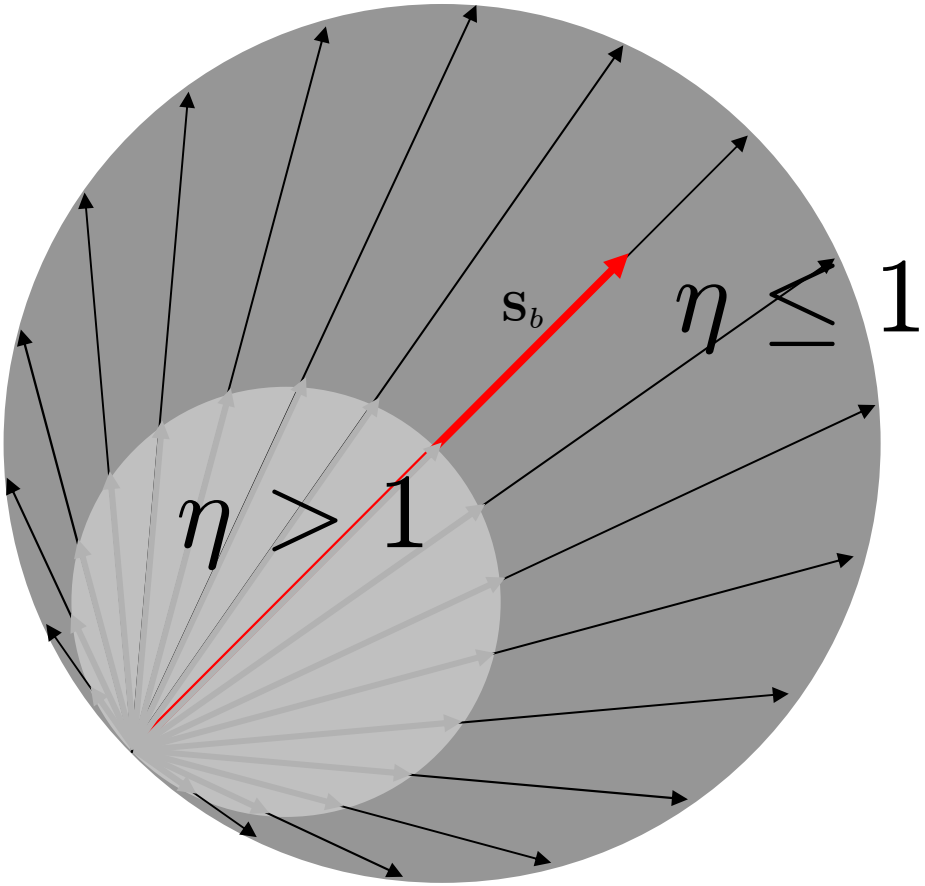
- ✓ Cross-space affinity matrix
  - ☞ We transform the affinity matrix $\mathbf{A}$ in another matrix $E \in R^{m,m}$ each element of which is defined as:

$$e_{i,j} = \frac{\langle \mathbf{S}_i, \mathbf{S}_j \rangle}{\|\mathbf{S}_i\|^2} = S(\pi_i, \pi_j)$$

  - ☞ Advantage: asymmetry of the chosen affinity kernel allows for a natural way to represent aggregations using directed graphs and find representative elements as those maximising the total semantic equivalence measured as $\sum_{k \neq j} e_{k,j}$

Geometric Properties

$$S(a,b) \geq \eta \Leftrightarrow \|\mathbf{S}_b\| \cos(\mathbf{S}_b, \mathbf{S}_a) \geq \eta \|\mathbf{S}_a\|$$



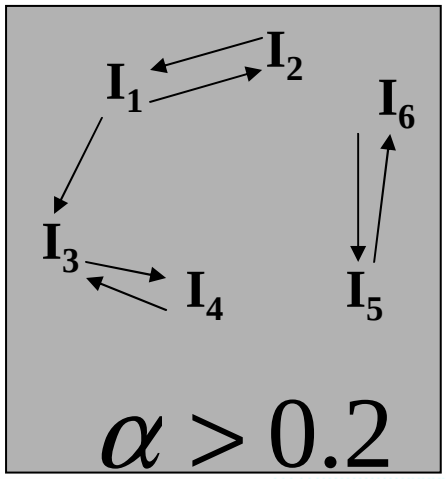
How does the aggregation work, and in what does generalisation consist?

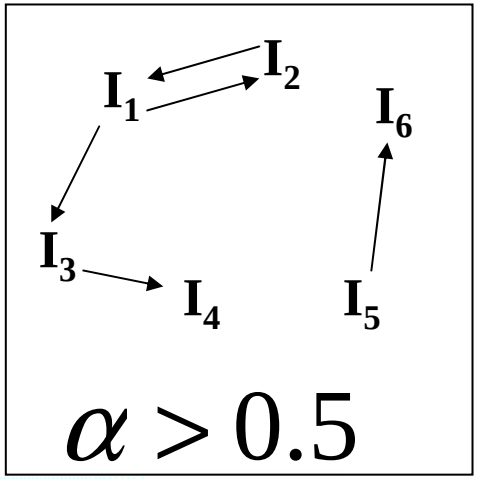
- ✓ The proposed schema can be interpreted as a generic aggregation mechanism between 2 sources of information elements streams
- ✓ The content of the *target* stream (A) is used as a ***dynamic dictionary*** of terms w.r.t. which the content of the *source* stream (B) is indexed
  - ☞ Dynamics stands in the fact that it is a stream…i.e. only the last N terms are considered
- ✓ Each aggregation x is made up of a subset of items of B which share common terms in A, as given by a semantic affinity measurement S
  - ☞ In our case it is asymmetric, but it could be the plain Cosine similarity
  - ☞ Semantic affinity is a generalisation of semantic similarity
- ✓ …**plus the terms of the dictionary A which are relevant to each aggregated element of x**
- ✓ Aggregations are generated through the visit of the semantic affinity matrix E

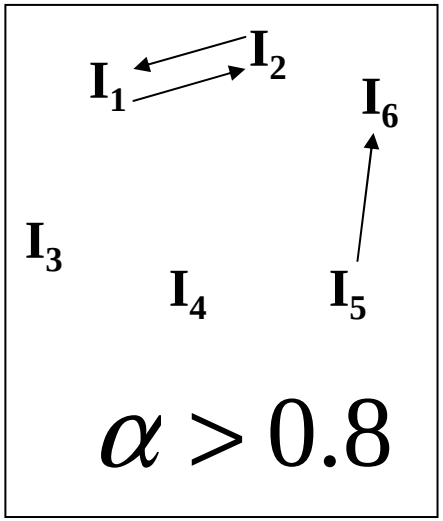
Core technologies: hybrid clustering

$E =$

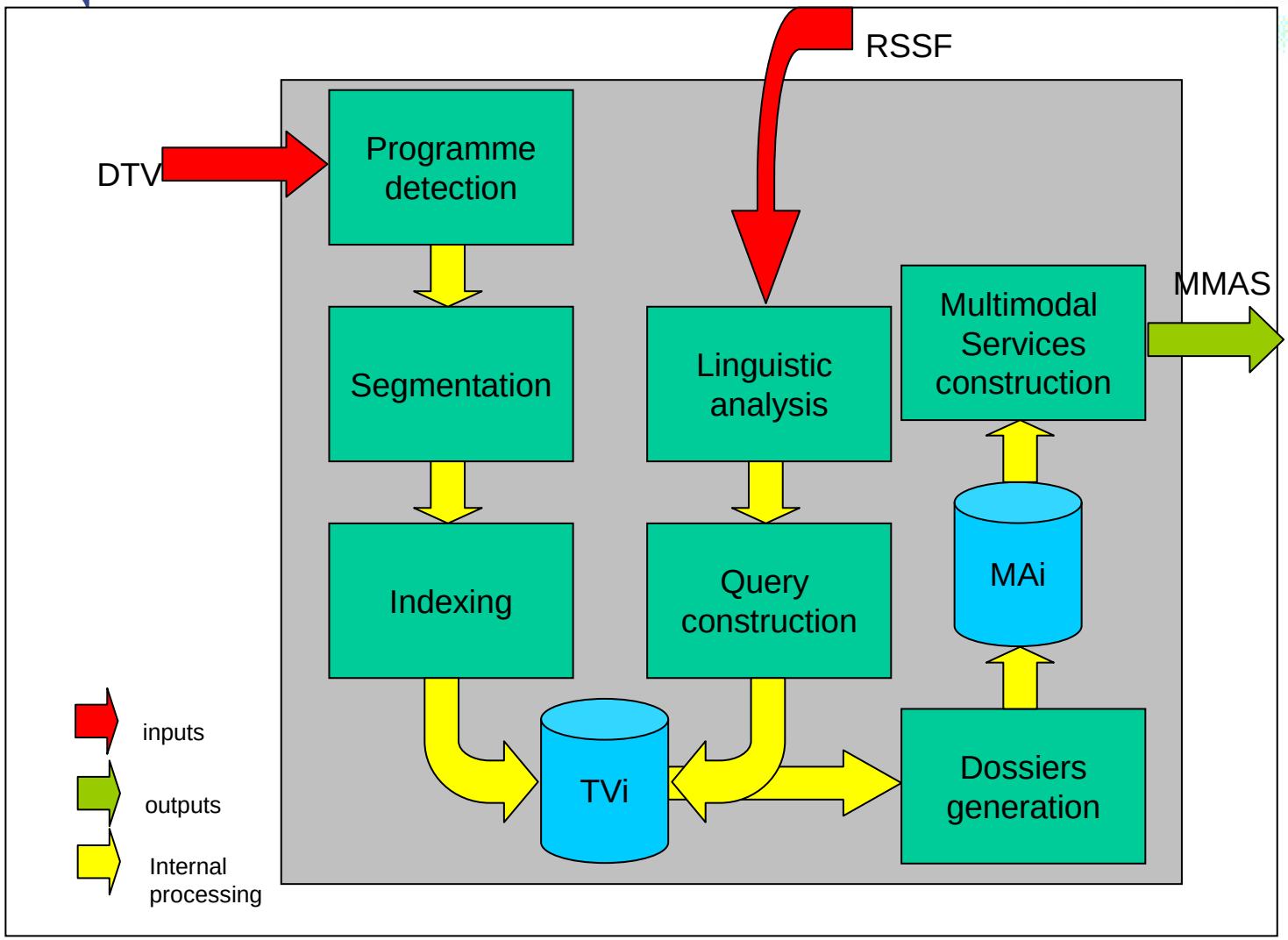
	I_1	I_2	I_3	I_4	I_5	I_6
I_1	1	0.82	0.53	0	0	0
I_2	0.85	1	0	0	0	0
I_3	0	0	1	0.6	0	0
I_4	0	0	0.30	1	0	0
I_5	0	0	0	0	1	0.9
I_6	0	0	0	0	0.28	1







Architecture



Core technologies: programme detection and segmentation

- ✓ Visual pattern matching algorithm for programme detection
  - ☞ Training on detection starting and ending jingles
  - ☞ Feature selection based of Bhattacharyya distance of feature distribution
- ✓ Three-layered heuristics framework for segmentation
  - ☞ H1: Most frequent speaker == anchorman
  - ☞ H2: Story boundaries are at anchorman appearances
  - ☞ H3: Shot changes during anchorman's speech are story boundaries

Implementation of the semantic affinity concept

- ✓ The semantic affinity function $s_{i,j} = R(\pi_i, \beta_j)$ has been implemented as the score of queries automatically built from the content of RSS items, run on the television items full text index.

$$R(\pi_i, \beta_j) = \max\{I(q(L(s(\pi_i))), t(\beta_j)) - \zeta, 0\}$$

Example of query construction

Violence in Kenya, 50 deads in a burned church.
Victims had found shelter from fights after presidential elections.
European Union asks for inquiry on votes.

← Original text

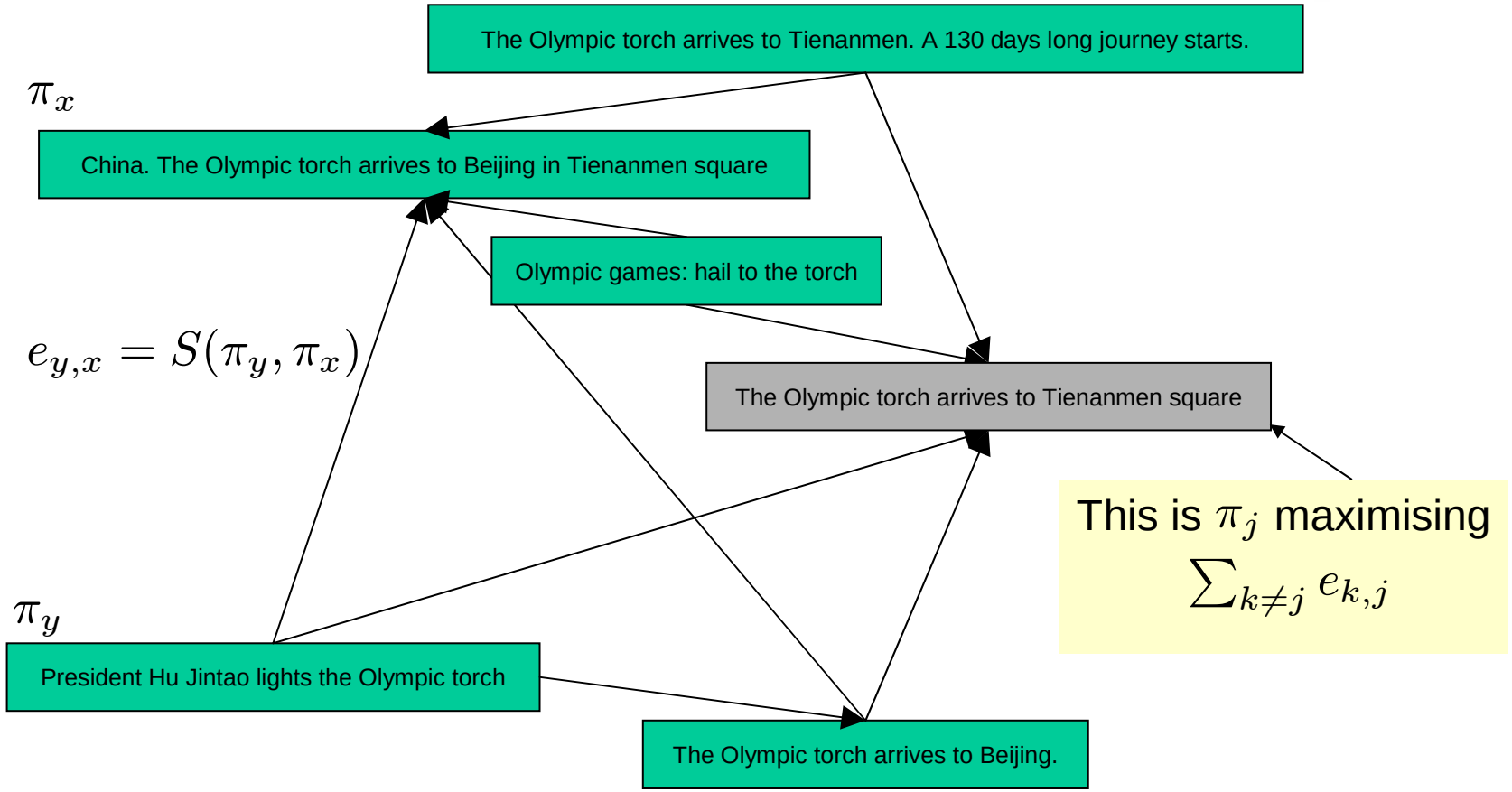
Violence in Kenya, 50 deads in a burned church.
Victims had found shelter from fights after presidential elections.
European Union asks for inquiry on votes.

← Tagged nouns

(Violence Kenya deads church)^4
OR
(victims shelter fights) ^2
OR
(European Union inquiry votes)

← Resulting query (Lucene syntax)

Graph example



Example of publication

Sottoscrizione al feed

Brasiliani Milan incontrano Lula (Aggiornato a 2008-11-12 09:28:52+01)

Title and discovery timestamp

Oggi 12 novembre 2008, 8 ore fa

Include contributi dai seguenti telegiornali:

canale5 - TG5 08:00 del 2008-11-11
rete4 - TG4 19:00 del 2008-11-11
rai3 - TG3 19:00 del 2008-11-11
canale5 - TG5 20:00 del 2008-11-11
rai1 - TG1 20:00 del 2008-11-11
rai2 - TG2 20:30 del 2008-11-11

Newscast items

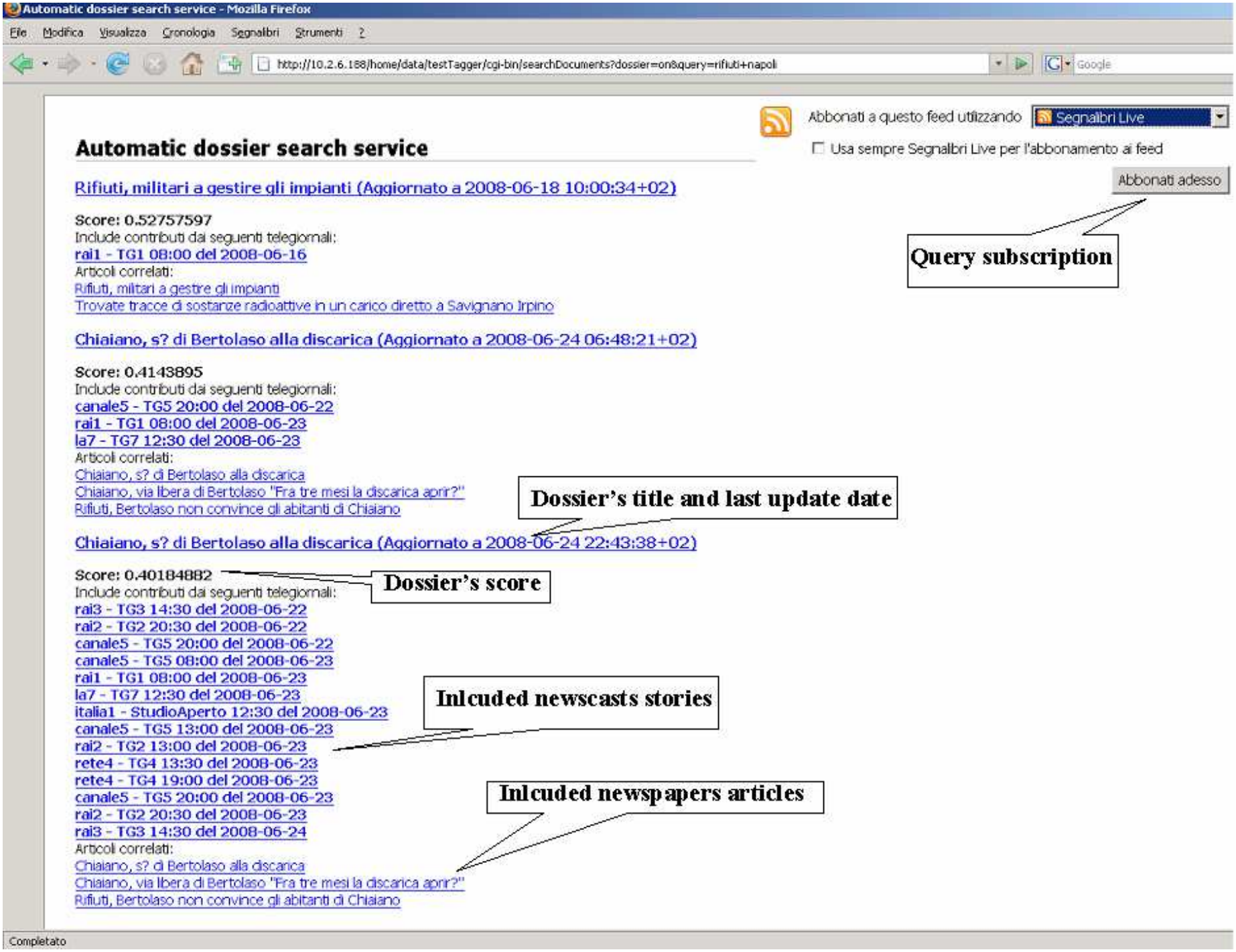
Articoli correlati:

http://www.tgcom.mediaset.it/mondo/articoli/articolo432794.shtml
http://www.repubblica.it/2008/11/sezioni/politica/berlusconi-varie/berlusconi-brasiliani/berlusconi-brasiliani.html?rss
http://www.ansa.it/site/notizie/awnplus/italia/news/2008-11-11_111299548.html
http://rss.feedsportal.com/c/32180/f/420290/s/259438d/l/0L0Sasca0Bit0Cmoddettnews0Bphp0Dcanale0FPOL0Gidnews0F790A0A710Garticolo0F190J3a0A90J20A0E0J20A
http://rss.feedsportal.com/c/32180/f/420290/s/259468c/l/0L0Sasca0Bit0Cmoddettnews0Bphp0Dcanale0FPOL0Gidnews0F790A0A760Garticolo0F190J3a220J20A0E0J20AIT
http://www.ansa.it/site/notizie/awnplus/mondo/news/2008-11-11_111303269.html
http://www.adnkronos.com/IGN/Esteri/?id=3.0.2700271812
http://www.ansa.it/site/notizie/awnplus/calcio/news/2008-11-11_111304296.html

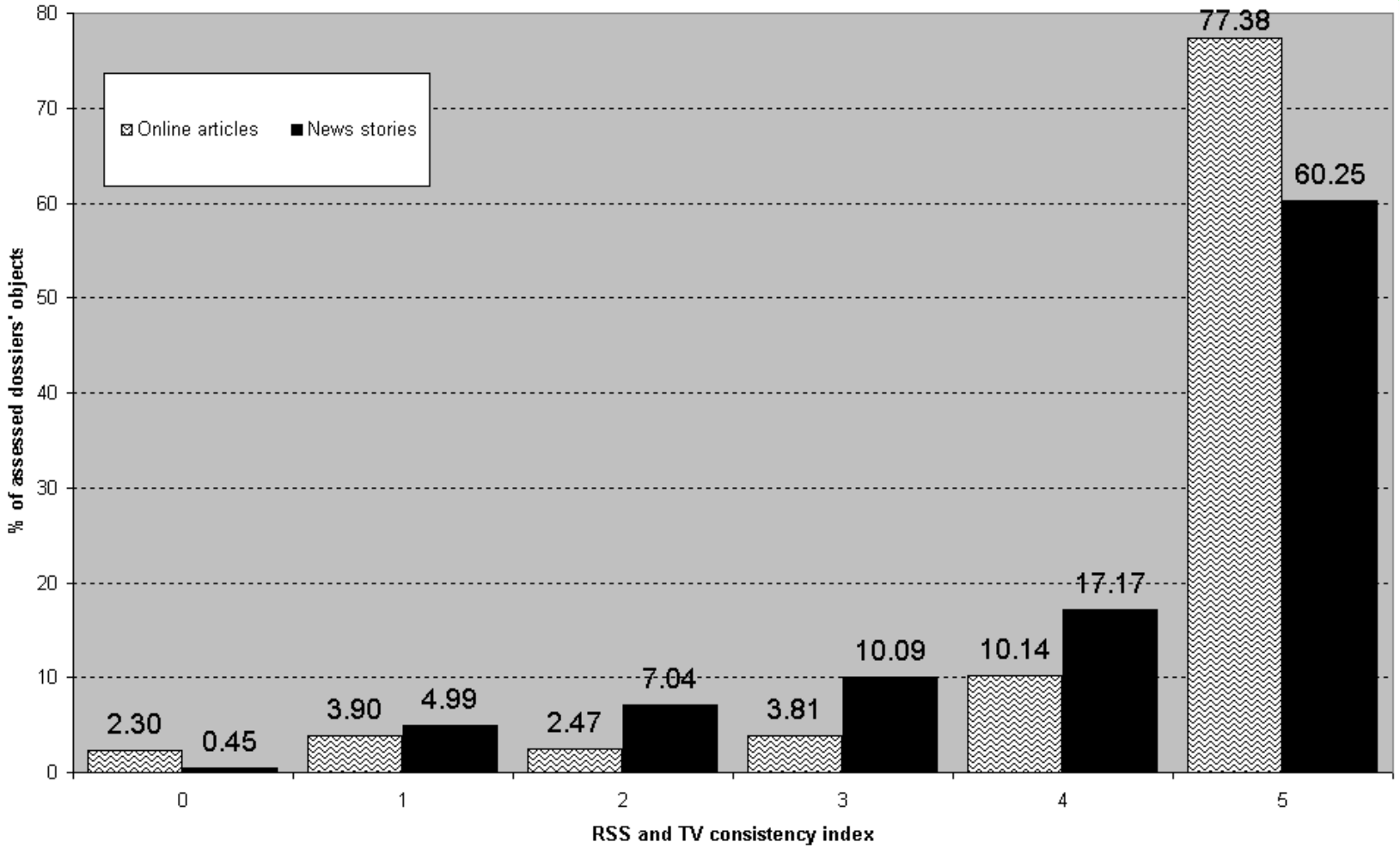
Web documents

Experimental evaluation

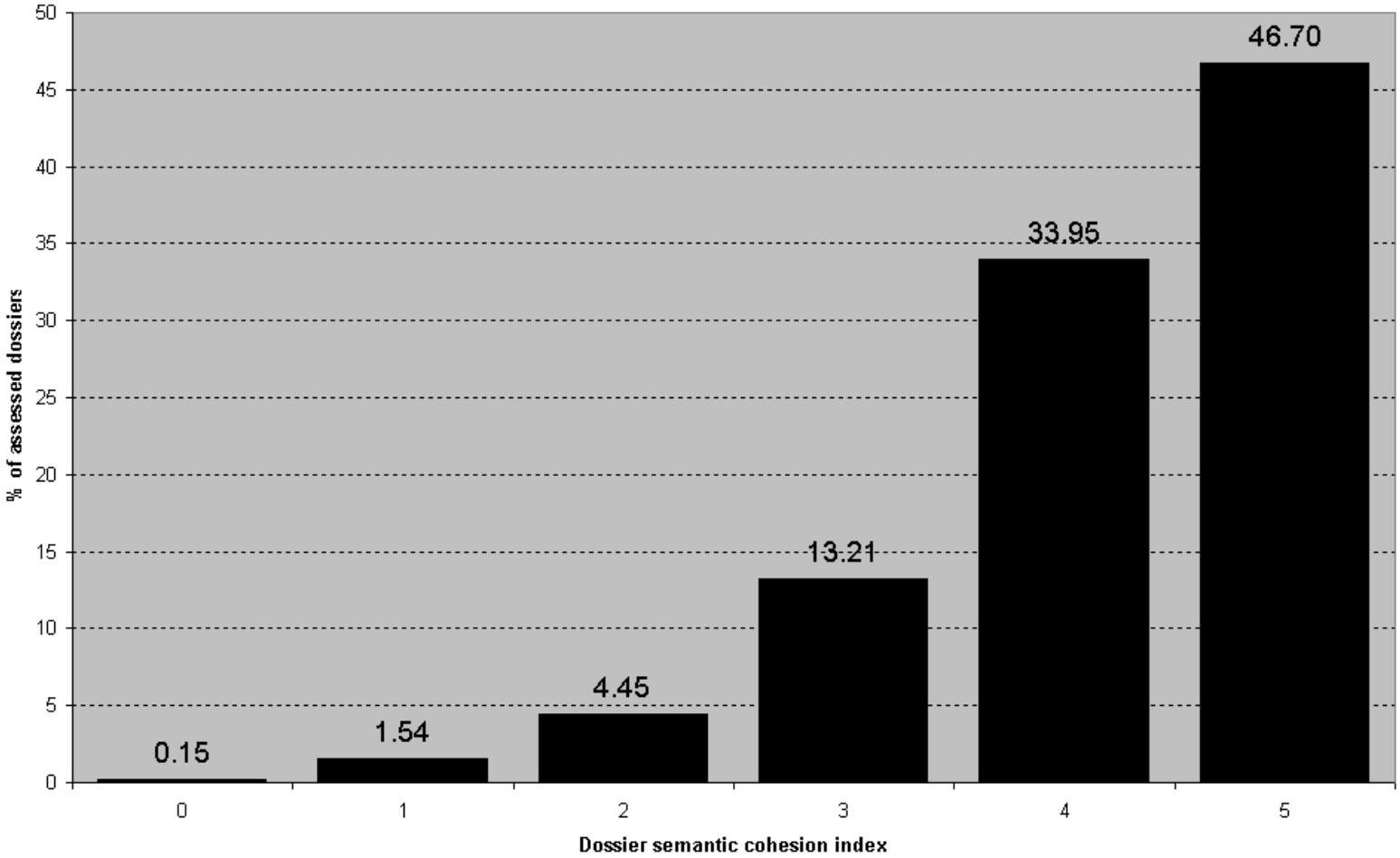
- ✓ Two main evaluations
  - ☞ Cross-modal aggregation system
    - ➢ Implementation of $s_{i,j}$ is done using the full text ranking of newscast items w.r.t. queries automatically constructed from the linguistic processing of RSS sentences
    - ➢ Intrinsic consistency of RSS items to the aggregation (I1)
    - ➢ Intrinsic consistency of newscast items to the aggregation (I2)
    - ➢ Overall semantic cohesion (I3)
    - ➢ Representativeness of the user-defined aggregation title (I4)
    - ➢ System-selected and user-defined title agreement (I5)
    - ➢ Representativeness of the correctly (I6) and wrongly (I7) system-selected titles
  - ☞ Search and retrieval system
    - ➢ Documents are collation of newscasts' items spoken text and RSS items sentences belonging to the same cross-modal aggregation
    - ➢ Mean Average Precision (MAP)
  - ☞ 25 expert users, 651 assessed aggregations


Automatic dossier search service - Mozilla Firefox
File Modifica Visualizza Cronologia Segnalibri Strumenti ?
http://10.2.6.188/home/data/testTagger/cgi-bin/searchDocuments?dossier=on&query=rifiuti+napoli
Abbonati a questo feed utilizzando Segnalibri Live
Usa sempre Segnalibri Live per l'abbonamento ai feed
Abbonati adesso
Automatic dossier search service
Rifiuti, militari a gestire gli impianti (Aggiornato a 2008-06-18 10:00:34+02)
Score: 0.52757597
Include contributi dai seguenti telegiornali:
rai1 - TG1 08:00 del 2008-06-16
Articoli correlati:
Rifiuti, militari a gestire gli impianti
Trovate tracce di sostanze radioattive in un carico diretto a Savignano Irpino
Chiaiano, s? di Bertolaso alla discarica (Aggiornato a 2008-06-24 06:48:21+02)
Score: 0.4143895
Include contributi dai seguenti telegiornali:
canale5 - TG5 20:00 del 2008-06-22
rai1 - TG1 08:00 del 2008-06-23
la7 - TG7 12:30 del 2008-06-23
Articoli correlati:
Chiaiano, s? di Bertolaso alla discarica
Chiaiano, via libera di Bertolaso "Fra tre mesi la discarica aprir?"
Rifiuti, Bertolaso non convince gli abitanti di Chiaiano
Chiaiano, s? di Bertolaso alla discarica (Aggiornato a 2008-06-24 22:43:38+02)
Score: 0.40184882
Include contributi dai seguenti telegiornali:
rai3 - TG3 14:30 del 2008-06-22
rai2 - TG2 20:30 del 2008-06-22
canale5 - TG5 20:00 del 2008-06-22
canale5 - TG5 08:00 del 2008-06-23
rai1 - TG1 08:00 del 2008-06-23
la7 - TG7 12:30 del 2008-06-23
italia1 - StudioAperto 12:30 del 2008-06-23
canale5 - TG5 13:00 del 2008-06-23
rai2 - TG2 13:00 del 2008-06-23
rete4 - TG4 13:30 del 2008-06-23
rete4 - TG4 19:00 del 2008-06-23
canale5 - TG5 20:00 del 2008-06-23
rai2 - TG2 20:30 del 2008-06-23
rai3 - TG3 14:30 del 2008-06-24
Articoli correlati:
Chiaiano, s? di Bertolaso alla discarica
Chiaiano, via libera di Bertolaso "Fra tre mesi la discarica aprir?"
Rifiuti, Bertolaso non convince gli abitanti di Chiaiano
Completato
Query subscription
Dossier's title and last update date
Dossier's score
Inlcuded newscasts stories
Inlcuded newspapers articles

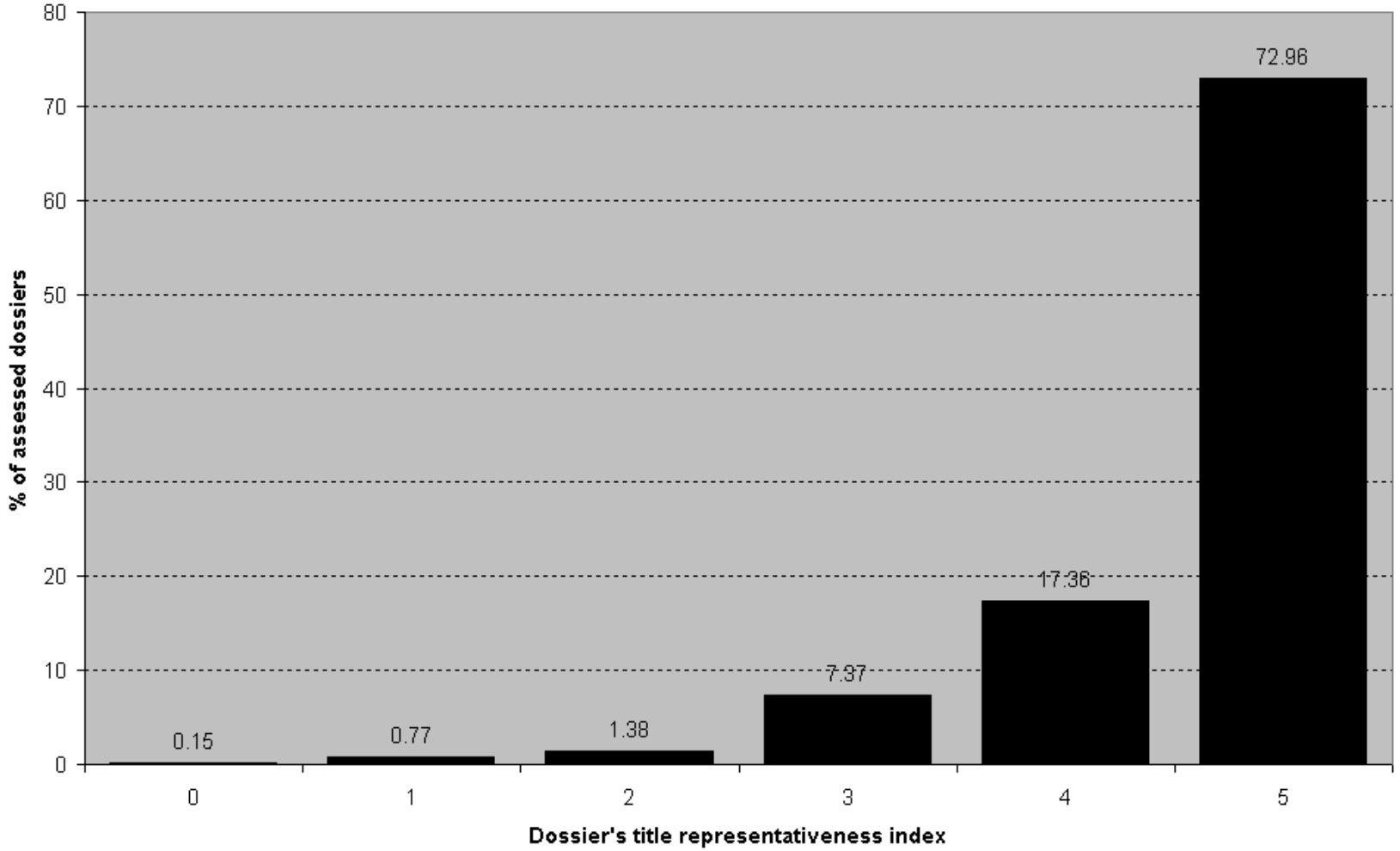
Results (I1, I2)



Results (I3)


50
45
40
35
30
25
20
15
10
5
0
% of assessed dossiers
0.15
1.54
4.45
13.21
33.95
46.70
0
1
2
3
4
5
Dossier semantic cohesion index

Results (I4)


80
70
60
50
40
30
20
10
0
% of assessed dossiers
0.15
0.77
1.38
7.37
17.36
72.96
0
1
2
3
4
5
Dossier's title representativeness index

Results (I5, I6, I7 Vs. I1 – I4)

Index	I1	I2	I3	I4	I5	I6	I7
Mean	4.23	4.65	4.24	4.66	1.55	4.85	4.63
Stdev.	0.85	0.89	1.17	0.70	-	0.62	0.77
Cinf s.	4.19	4.62	4.20	4.59	-	4.84	4.47
Cinf e.	4.35	4.68	4.28	4.70	-	4.89	4.69

Mean → Average score

Stdev. → Standard Deviation

Cinf s. → Confidence Interval Start

Cinf e. → Confidence Interval End

Results(MAP)

- ✓ We used the same users panel to evaluate a set of 50 queries made to the system
- ✓ MAP=0.79

Conclusions

- ✓ The developed cross-modal aggregation system is based on a generic hybrid clustering scheme
- ✓ Experimental results are very encouraging
- ✓ Future works will regard both theoretical and implementation aspects
  - ☞ Better exploitation of graphs information
  - ☞ Extension to more than 2 aggregated sources

a.messina@rai.it; maurizio.montagnuolo@rai.it